

Liu, G., MacLean, C., MacDonald, P., Major, J., Matthews, C., McCarthy, M., Melarini, I., Meneus, I., Mihov, T., Mlengi, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, P., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunh, P., Pierre, N., Raymond, C., Petta, F., Pise, C., Porev, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Romann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, B., Viel, R., Vo, A., Wilson, R., Wu, X., Wyman, D., Young, G., Zaidoun, J., Zemek, L., Zimmer, A., and Zody, M.

TITLE JOURNAL

Submitted (06-Aug-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 27, 2001 this sequence version replaced j1.16356513.
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L838
Center clone name: 81_A_22

----- Summary Statistics
Sequencing vector: M13: M77815; 40% of reads
Sequencing vector: Plasmid; n/a; 60% of reads
Chemistry: Dye-primer-amersham; 2% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 168221 bases at least Q40
Consensus quality: 175650 bases at least Q30
Consensus quality: 178768 bases at least Q20
Insert size: 203600; agarose-1p

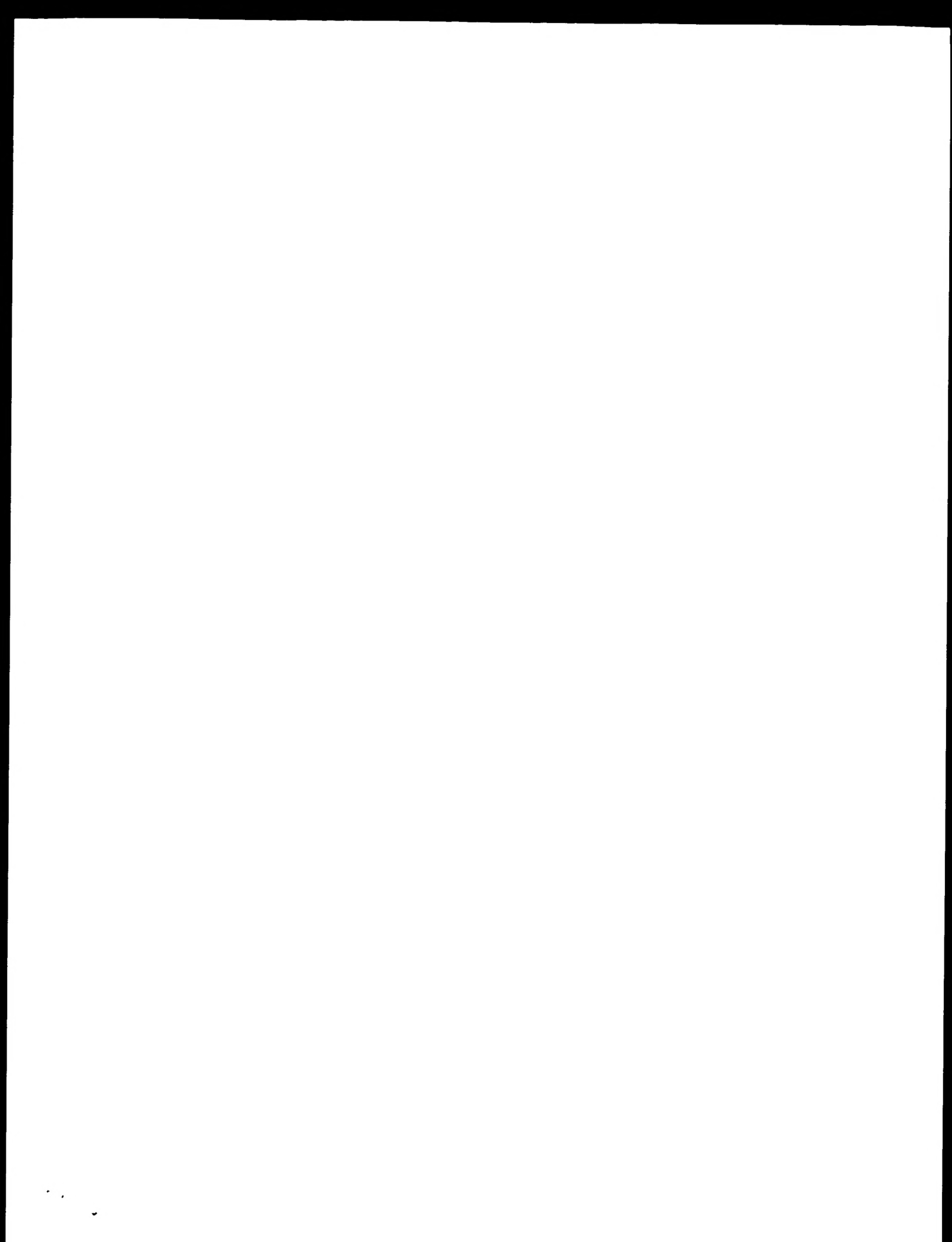
Insert size: 181495; sum-of-contigs
Quality coverage: 7.6 in Q20 bases.

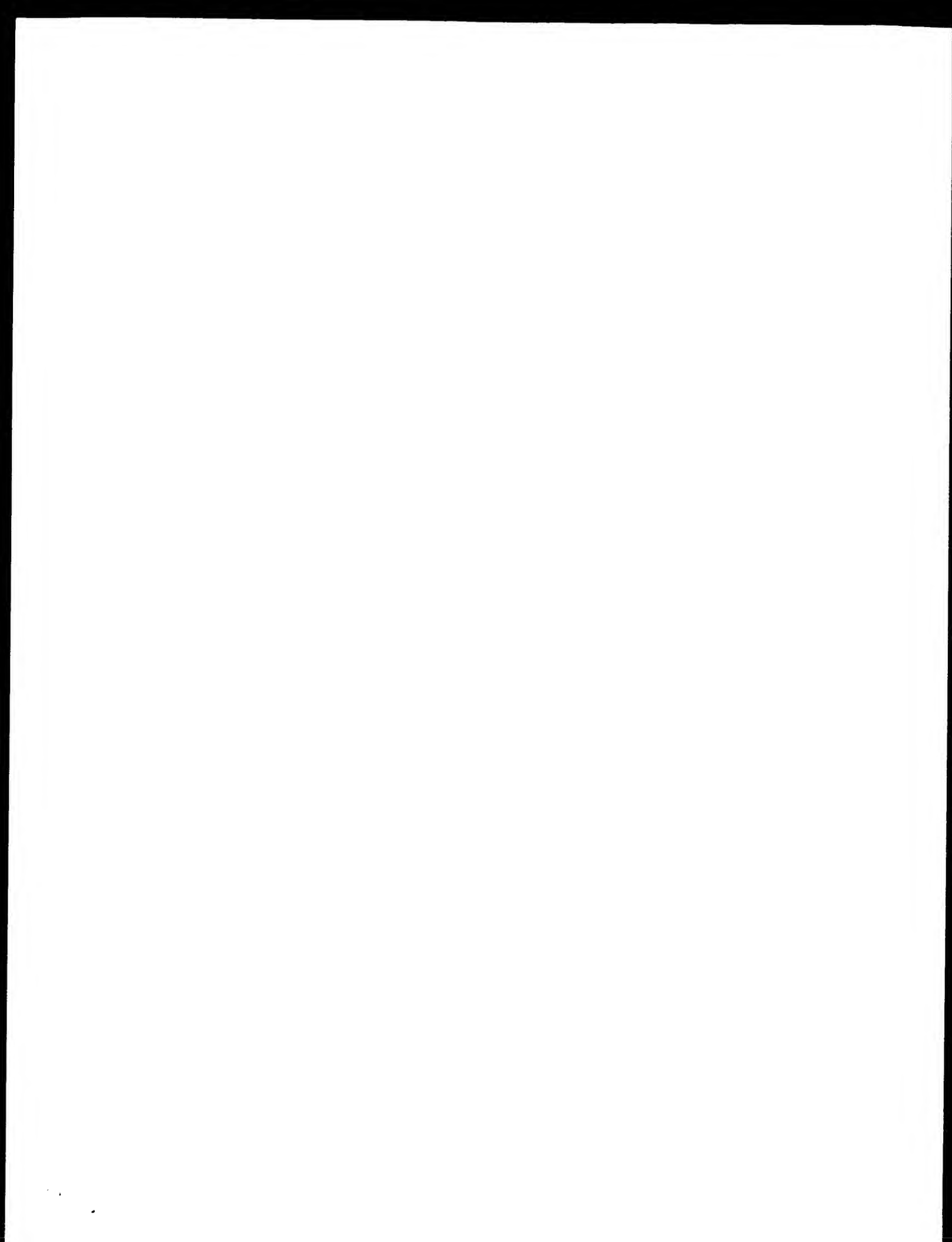
* NOTE: This is a 'working draft' sequence. It currently consists of 32 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 10663: contig of 10663 bp in length
* 10664 10763: gap of 100 bp
* 10764 11392: contig of 629 bp in length
* 11393 11492: gap of 100 bp
* 11493 12333: contig of 841 bp in length
* 12334 12433: gap of 100 bp
* 12434 13332: contig of 899 bp in length
* 13333 13432: gap of 100 bp
* 13433 14716: contig of 1284 bp in length
* 14717 14816: gap of 100 bp
* 14817 16101: contig of 1285 bp in length
* 16102 16201: gap of 100 bp
* 16202 17797: contig of 1596 bp in length
* 17798 17897: gap of 100 bp
* 17898 18949: contig of 1052 bp in length
* 18950 19049: gap of 100 bp
* 19050 20494: contig of 1445 bp in length
* 20495 20594: gap of 100 bp
* 20595 21775: contig of 1181 bp in length
* 21776 21875: gap of 100 bp
* 21876 23459: contig of 1584 bp in length
* 23460 23559: gap of 100 bp
* 23560 25811: contig of 2252 bp in length
* 25832 25931: gap of 100 bp
* 25932 27763: contig of 1832 bp in length
* 27764 27863: gap of 100 bp
* 27864 29086: contig of 2123 bp in length
* 29087 30086: gap of 100 bp

* 30087 32675: contig of 2589 bp in length
* 32676 32775: gap of 100 bp
* 32776 46266: contig of 3491 bp in length
* 46267 46366: gap of 100 bp
* 46367 49678: contig of 3321 bp in length
* 49688 39787: gap of 100 bp
* 39788 42811: contig of 3024 bp in length
* 42812 42911: gap of 100 bp
* 42912 47748: contig of 4837 bp in length
* 47749 47848: gap of 100 bp
* 47849 52056: contig of 4208 bp in length
* 52057 52156: gap of 100 bp
* 52157 56257: contig of 4101 bp in length
* 56258 56357: gap of 100 bp
* 56358 62535: contig of 6178 bp in length
* 62536 62635: gap of 100 bp
* 62636 65138: contig of 2503 bp in length
* 65139 65238: gap of 100 bp
* 65239 65640: contig of 4402 bp in length
* 65641 69740: gap of 100 bp
* 69741 74838: contig of 5098 bp in length
* 74839 74938: gap of 100 bp
* 74939 79482: contig of 4544 bp in length
* 79483 79582: gap of 100 bp
* 79583 81688: contig of 2306 bp in length
* 81689 81988: gap of 100 bp
* 81989 89859: contig of 7871 bp in length
* 89860 89959: gap of 100 bp
* 89960 158380: contig of 58421 bp in length
* 158381 158480: gap of 100 bp
* 158481 165923: contig of 8443 bp in length
* 165924 167923: gap of 100 bp
* 167924 176904: contig of 9881 bp in length
* 176905 177004: gap of 100 bp
* 177005 184595: contig of 7591 bp in length.

FEATURES
Source
1. 184595
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chr="chromosome-17"
/map="17"
/clone="RP11-81A22"
/clone_lib="RP11-11 Human Male BA"
1. 10663
/note="assembly_fragment"
clone_end:Sp6
vector_side:left
10764..11392
/note="assembly_fragment"
11493..12333
/note="assembly_fragment"
12434..13332
/note="assembly_fragment"
13433..14716
/note="assembly_fragment"
14817..16101
/note="assembly_fragment"
16202..17797
/note="assembly_fragment"
17898..18949
/note="assembly_fragment"
19050..20494
/note="assembly_fragment"
20595..21775
/note="assembly_fragment"
21876..23459
/note="assembly_fragment"
23560..25831
/note="assembly_fragment"
25932..27763
/note="assembly_fragment"
27864..29086
/note="assembly_fragment"





us-09-727-770-2.p2n.rng

Tue Dec 31 14:03:15 2002

```

Pred. No.: 5,496-106 Length: 612
Score: 1046.00 Matches: 203
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.05% Indels: 0
DB: 22 Gaps: 0

US-09-727-770-2 (1-205) x ABA64433 (1-612)
QY 1 MetSerAsnSerProGlyThrAlaLeuValPheThrIleSerGlyAlaMetAlaThr 20
DB 3 ATGTGCAACAAAGCCGACGATGCTTGGTTTTCACCAATCGGGTGTATGCGCACC 62
QY 21 MetValSerSerGlyLeuAlaAlaGlyGlyMetAlaIleAsnGlyThrGlyIleMet 40
DB 63 ATGTGCTCCAGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 122
QY 41 AlaMetSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValIleMet 60
DB 123 GCGATGCTGTGATGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 182
QY 61 AlaGlyIleIleThrIleThrGlyLeuValAlaAlaValProProAlaAsnSerLeuAsn 80
DB 183 GCTGTGATATCATCACCATCATGCGCTAGTGGGGGTGTGGGGGTGTGGGGGTGTGG 242
QY 81 AspAspAsnSerLeuIleYrSerSerPheLeuGluLeuGlyAlaGlyLeuSerGlyLeuAla 100
DB 243 GATGCAACAACTCTATATACAGATTTCTTCCAGCTGGGCTGTGGCTGTGGCTGTGG 302
QY 101 AlaGlyPheAlaIleValIleValGlyAspThrGlyIleGlySerGlyIleThrPro 120
DB 403 GCGGCTCTTGGCAGTCATCTGTCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 362
QY 121 ArdeupPheValGlyMetIleLeuIleLeuIlePheAlaLysValLeuIleLeuSerThr 140
DB 363 GCACTATTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 422
QY 141 LysGluProLeuSerLysProThrSerIleArgIleArGlyLysAspHisProSerSer 160
DB 423 AAGCAGCGCTCTCTCAAAAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 482
QY 163 PheArGlyLysGluProAspThrHisValIleGlySerTrpProSerValIleAspLeu 180
DB 483 TTCGCAACAAACAGCCGTCAGACGATGCTGGGCGAGCTGGGCGAGCTGGGCGAGC 542
QY 181 LeuSerValGluGlyProArValIleIleSerLeuAlaAlaProGlyProLeuProPro 200
DB 543 CTAAGTGTACAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
QY 201 HisAlaVal 203
DB 603 CATGCTGTG 611

RESULT 4
ID ABA64433
XX ABA64433 standard: DNA: 612 BP.
AC ABA64433:
XX ABA64433:
XX ABA64433:
DI 01-FEB-2002 (first entry)
XX Human foetal liver single exon nucleic acid probe #12738.
DE Human foetal liver; gene expression; single exon nucleic acid probe; ss.
KW Homo sapiens.
XX Homo sapiens.
XX W-20015277-A2
XX W-20015277-A2
XX 09-AUG-2001.
XX 30-JAN-2001; 2001W0-0500669.
XX

```

```

PR 04-FEB-2000; 2000US-0180412.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0642466.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GR-0024263.
XX
XX (MODE ) MOLECULAR DYNAMICS INC.
XX From SG, Hanzel JK, Chen W, Rank DK
XX WP1: 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver.
XX
XX Claim 4: SEQ ID NO 12738; 63pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at http://wipo.int/pat/publicated\_pat\_sequences.
XX
XX Sequence 612 BP; 120 A; 195 C; 151 G; 146 T; 0 other.

```

```

Alignment Scores:
Pred. No.: 5,496-106 Length: 612
Score: 1046.00 Matches: 203
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.05% Indels: 0
DB: 22 Gaps: 0

```

```

US-09-727-770-2 (1-205) x ABA64433 (1-612)
QY 1 MetSerAsnSerProGlyThrAlaLeuValPheThrIleSerGlyAlaMetAlaThr 20
DB 3 ATGTGCAACAAAGCCGACGATGCTTGGTTTTCACCAATCGGGTGTATGCGCACC 62
QY 21 MetValSerSerGlyLeuAlaAlaGlyGlyMetAlaIleAsnGlyThrGlyIleMet 40
DB 63 ATGTGCTCCAGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 122
QY 41 AlaMetSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValIleMet 60
DB 123 GCGATGCTGTGATGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 182
QY 61 AlaGlyIleIleThrIleThrGlyLeuValAlaAlaValProProAlaAsnSerLeuAsn 80
DB 183 GCTGTGATATCATCACCATCATGCGCTAGTGGGGGTGTGGGGGTGTGGGGGTGTGG 242
QY 81 AspAspAsnSerLeuIleYrSerSerPheLeuGluLeuGlyAlaGlyLeuSerGlyLeuAla 100
DB 243 GATGCAACAACTCTATATACAGATTTCTTCCAGCTGGGCTGTGGCTGTGGCTGTGG 302
QY 101 AlaGlyPheAlaIleValIleValGlyAspThrGlyIleGlySerGlyIleThrPro 120
DB 403 GCGGCTCTTGGCAGTCATCTGTCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 362
QY 121 ArdeupPheValGlyMetIleLeuIleLeuIlePheAlaLysValLeuIleLeuSerThr 140
DB 363 GCACTATTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 422
QY 141 LysGluProLeuSerLysProThrSerIleArgIleArGlyLysAspHisProSerSer 160
DB 423 AAGCAGCGCTCTCTCAAAAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 482

```


Tue Dec 31 14:03:15 2002

cy and hyaline membrane disease. The present sequence is a single exon probe open reading frame of this invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from Wipo at http://wipo.int/pdb/pdb-abstract-f_333-press

XX Sequence 612 BP; 120 A; 195 C; 151 G; 146 T; 0 other;

Alignment Scores:

Prod. No.:	5,396-106	Length:	612
Score:	1046.00	Matches:	203
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.05%	Indels:	0
DB:	24	Gaps:	0

US-09-727-770-2 (1-205) x ABA9210 (1-612)

QY 1 MetSerAsnSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThr 20

Db 3 ATCTCAACACAGGCGGACGATATATGGCTTTTCACCACTGCGGTCTAAGCGAC 62

QY 2 MetValSerSerGlyLeuGlyAlaAlaPheGlyMetAlaGlyAsnGlyThrGlyIleMet 40

Db 63 AIGCTCCAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 122

QY 41 AlaMetSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValValMet 60

Db 123 GCGATGTCGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 182

QY 61 AlaglyIleIleIleThrIleGlyLeuValAlaAlaValProProAlaAsnSerLeuAsn 80

Db 183 GCTGGTATCAACCATCTATGCGCTAGTGGCGGCTGCTGGCTGGCTGGCTGGCTGG 242

QY 81 AspAspAsnSerLeuTyrSerSerPheLeuGlyIleGlyAlaGlyLeuSerGlyIleAla 100

Db 243 GATGCAACACATCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302

QY 101 AlaglyIleAlaIleValIleValGlyAspGlyGlyIleGlyIleGlyIleGlyIle 120

Db 303 GCGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362

QY 121 AraLeuPheValGlyMetIleLeuIleLeuIlePheAlaGlyValLeuIleLeuSerThr 140

Db 363 GCAATATTGTAGCATGATCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422

QY 141 LysGlnProLeuSerLysProThrSerHisAlaIleAlaGlySerGlyIleGlySer 160

Db 424 AAGTAGCGCTCTCAAAACGACGACGACGACGACGACGACGACGACGACGACGAC 482

QY 161 PheArgAsnLysGlnProAspThrHisValIleGlySerTrpProSerValIleAspLeu 180

Db 484 TTCGGACAAACACACGCTGACGACGACGACGACGACGACGACGACGACGACGAC 542

QY 181 LeuSerValGlnCysProArgValHisArgLeuLeuAlaArgProCysProLeuProPro 200

Db 544 CTAAGTGTACAGTGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602

QY 201 HisAlaVal 204

Db 603 CATGTGTGTG 611

RESULT 12

ABA9210

XX ABA9210 standard; DNA: 1160 BP.

XX ABA9210;

XX 29-JUL-2002 (first entry)

XX Human tumour specific anti-ischemic peptide encoding sequence #2.

XX

Proton pump inhibitor; tumour specific anti-ischemic peptide; cytosolic; tumour; ds; human.

Hom sapiens.

Key Location/Qualifiers

229..696

/ftag: a

/product "tumour specific anti-ischemic peptide"

JP2001286284 A.

16-OCT-2001.

05-APR-2000; 2000JP-0103966.

05-APR-2000; 2000JP-0103966.

(SADU) SATO N.

(SADU) SATO N.

(YAMA) YAMAGUCHI M.

WI: 2002-144186/18.

P-PSDB: ABB08456.

Gene diagnostic agent and gene treating agent for tumours comprises using tumour-specific antigen and proton pump inhibitor as antigen agent

claim 9; Page 49-40; 80pp; Japanese.

This invention relates to a diagnostic agent or treating agent for tumours by using a tumour specific antigen and a proton pump inhibitor. The method is cytostatic in its action and the tumour-specific antigen peptide is useful for treating and preventing cancers. This nucleotide sequence is a human tumour-specific antigenic peptide-encoding sequence.

Sequence 1160 BP; 175 A; 429 C; 412 G; 253 T; 0 other;

Alignment Scores:

Prod. No.:	2,596-59	Length:	1160
Score:	626.50	Matches:	149
Percent Similarity:	74.27%	Conservative:	10
Best Local Similarity:	68.66%	Mismatches:	44
Query Match:	59.33%	Indels:	16
DB:	24	Gaps:	4

US 09 727 770 2 (1-295) x ABA9210 (1-1160)

QY 2 SerAsnAsnSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThrMet 2

Db 238 TCCAAACAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 292

QY 22 ValSerSerGlyLeuGlyAlaAlaPheGlyMetAlaGlyAsnGlyThrGlyIleMetAla 41

Db 298 GTCCTCAGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 457

QY 42 MetSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValValMetAla 61

Db 458 ATGCTGCTATGCGGCGGACGACGACGACGACGACGACGACGACGACGACGAC 414

QY 62 GlyIleIleThrIleGlyGlyLeuValAlaValProProAlaAsnSerLeuAsnAsp 81

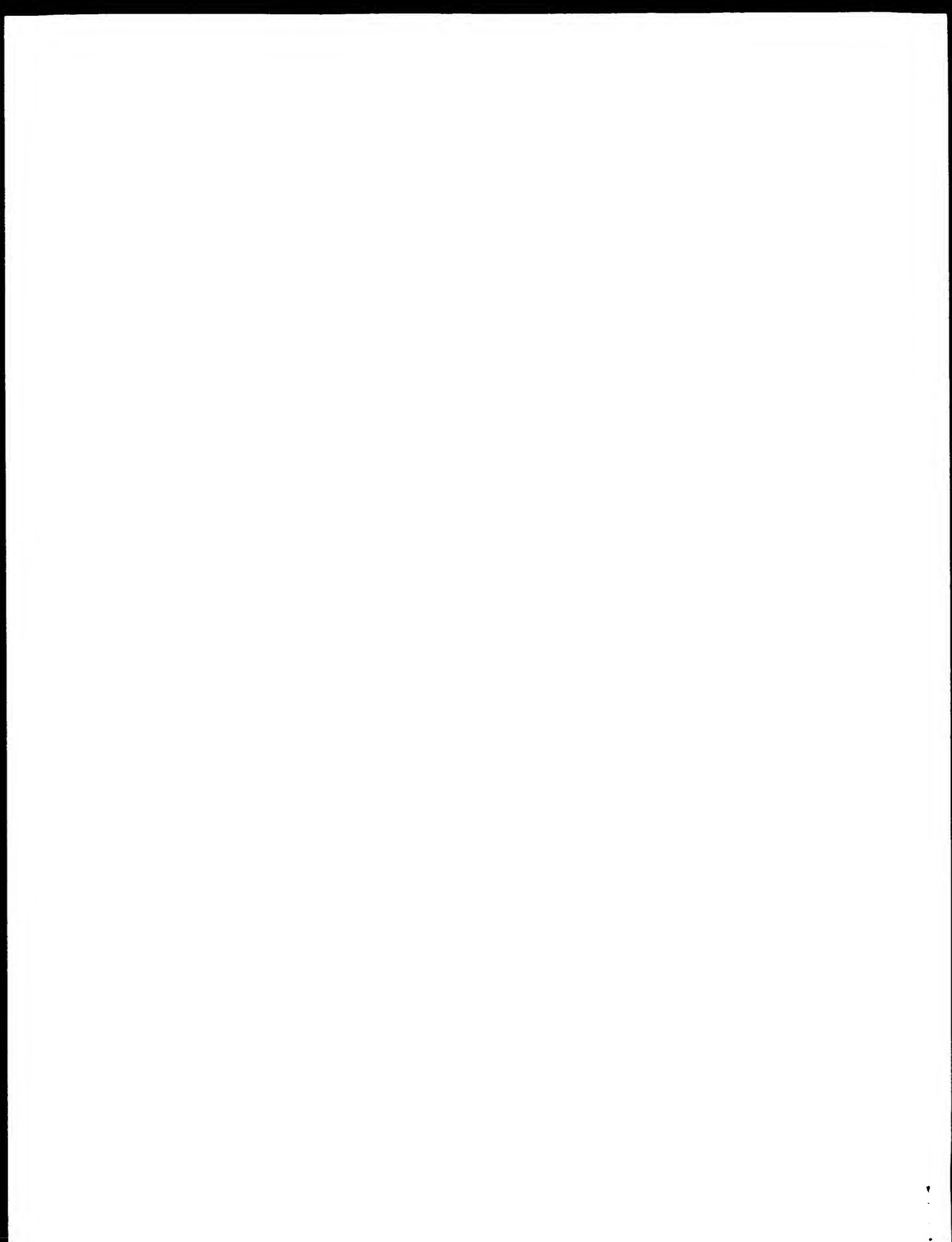
Db 415 GGTATCATCGCATCTAGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 474

QY 62 AspAsnSerLeuTyrSerSerPheLeuGlyIleGlyAla 97

Db 475 GACATCATCGCTTACCAACAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 534

QY 98 GlyLeuAlaAlaGlyPheAlaIleValIleValIleAspThrIleGlySerGlyThrAla 117

Db 535 GCGCTGGGAGCGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594



Percent Similarity: 100.00% Conservation: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-727-770-2 (1-205) x US-09-727-770-1 (1-639)

QY 1 MetSerAsnAspSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThr 20
 DB 10 ATGTCACCAACACAGCCGAGTATGCTTTGCTTTTCACATCTGAGGCTGATGCGGAGC 69
 QY 21 MetValSerSerGlyLeuGlyAlaAlaPheGlyMetAlaTyrAsnSerGlyThrGlyIleMet 40
 DB 130 GCGATGCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 189
 QY 61 AlaMetSerValMetTyrProGluLeuIleHisSerThrIleHisSerThrIleHisSer 80
 DB 190 GCTGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 249
 QY 81 AspAspAsnSerLeuTyrSerSerPheLeuGluLeuGlyAlaGlyLeuSerGlyLeuAla 100
 DB 250 CATCACAAACATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 309
 QY 101 AlaGlyPheAlaIleValIleValIleValIleValIleValIleValIleValIleValIle 120
 DB 410 GCGGCTTTTGGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 369
 QY 121 ArgLeuPheValGlyMetIleLeuIleLeuIleLeuIleLeuIleLeuIleLeuIleLeuIle 140
 DB 370 GCGATATTGTAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 429
 QY 141 LysGluProLeuSerLysProThrSerHisArgIleAlaGlyCysLysAspIleProSerSer 160
 DB 440 AAGCAATGCTGCT 489
 QY 161 PheArgAsnLysGlyProAspThrHisValIleGlySerIleProSerValIleAspLeu 180
 DB 490 TTTCGGCAACAAACAGCCGTCACAGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 549
 QY 181 LeuSerValGlnCysProArgValHisArgLeuLeuAlaArgProCysProLeuProPro 200
 DB 550 CTAAGGTACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609

QY 201 HisAlaValAspIle 205

DB 610 CATGCTGTGGAGATC 624

RESULT 2

US-09-727-770-3

Sequence 3, Application: US/09727770

Patent No. US20020147305A1

GENERAL INFORMATION:

APPLICANT: WEL, Ming-Hui

APPLICANT: KETCHUM, Karen A.

APPLICANT: BEASLEY, Ellen M.

TITLE OF INVENTION: ISOLATED HUMAN TRANSFERRED PROTEINS

TITLE OF INVENTION: AND USES THEREOF

FILE REFERENCE: CLO000651

CURRENT APPLICATION NUMBER: US/09727770

CURRENT FILING DATE: 2000-12-04

PRIOR APPLICATION NUMBER: US 60/209,846

PRIOR FILING DATE: 2000-06-02

NUMBER OF SEQ ID NOS: 3

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 4

LENGTH: 639

TYPE: DNA

ORGANISM: HUMAN

US-09-727-770-3

Alignment Scores:

Prod. No.: 2,740-110

Score: 1056.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 10

Length: 639

Mismatches: 205

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-727-770-2 (1-205) x US-09-727-770-3 (1-639)

QY 1 MetSerAsnAspSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThr 20
 DB 3000 ATGTCACCAACACAGCCGAGTATGCTTTGCTTTTCACATCTGAGGCTGATGCGGAGC 4059
 QY 21 MetValSerSerGlyLeuGlyAlaAlaPheGlyMetAlaTyrAsnSerGlyThrGlyIleMet 40
 DB 3000 ATGTCACCAACACAGCCGAGTATGCTTTGCTTTTCACATCTGAGGCTGATGCGGAGC 4119
 QY 41 AlaMetSerValMetTyrProGluLeuIleHisSerThrIleHisSerThrIleHisSer 60
 DB 3120 GCGATGCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 4179
 QY 61 AlaGlyPheAlaIleValIleValIleValIleValIleValIleValIleValIleValIle 80
 DB 3180 GCGGCTTTTGGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 4239
 QY 81 AspAspAsnSerLeuTyrSerSerPheLeuGluLeuGlyAlaGlyLeuSerGlyLeuAla 100
 DB 3240 GCGGCTTTTGGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 4299
 QY 101 AlaGlyPheAlaIleValIleValIleValIleValIleValIleValIleValIleValIle 120
 DB 3300 GCGGCTTTTGGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 4359
 QY 121 ArgLeuPheValGlyMetIleLeuIleLeuIleLeuIleLeuIleLeuIleLeuIleLeuIle 140
 DB 3360 GCGATATTGTAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4419
 QY 141 LysGluProLeuSerLysProThrSerHisArgIleAlaGlyCysLysAspIleProSerSer 160
 DB 3420 AAGCAATGCTGCT 4479
 QY 161 PheArgAsnLysGlyProAspThrHisValIleGlySerIleProSerValIleAspLeu 180
 DB 3480 TTTCGGCAACAAACAGCCGTCACAGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 4539
 QY 181 LeuSerValGlnCysProArgValHisArgLeuLeuAlaArgProCysProLeuProPro 200
 DB 3540 CTAAGGTACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4599

RESULT 4

US-09-864-761-16890

Sequence 16890, Application: US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Bank, David R.

APPLICANT: Beasley, Ellen M.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GERM LINE DERIVED SINGLE EXON NUCLEOTIDE ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Accell-X-1

CURRENT APPLICATION NUMBER: US/09864761

CURRENT FILING DATE: 2001-05-24

PRIOR APPLICATION NUMBER: US 60/2190,412

PRIOR FILING DATE: 2000-02-04

Senior

REFERENCE: AERONICA-X-1
CURRENT APPLICATION NUMBER: US/09/804,761
CURRENT FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/186,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 63/207,450
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/832,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/326,259
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668

Downloaded from <http://ajphaphapublications.sagepub.com/> at 11:06 11 November 2014


```

Db 1305 AATAAAGGAAGAAATCAATTCAGTCTAAGCCCTGTTATATACATTTTC 1364
QY 115 -----Gly- ThrAlaGlnGlnProArgLeuPheValGI 125
Db 1365 AATGCAAAATTTGTGTTGATGACAGGCAATGCTCAACAGCAAAAGCTTTTCGTGG 1424
QY 125 yMetIleIeuIleLeuIlePheAlaIzSValIleuIleIeu 138
Db 1425 TATGATTCTTAACCCATTTTGGCGAAGGCTCTTGCTCTA 1464

```

Search completed: December 29, 2002, 12:21:56
 Job time : 60 secs

GenCore version 5.1.4
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 28, 2002, 10:26:51 : Search time 2575 seconds
(without alignments)
2310,922 Kbits, 3.11 updates/sec

File: us-09-727-770-2
Perfect score: 1056
Sequence: 1 MSNNSPEYALVPIISGAMAI.....PVHRLIARCPHPHRAVDI 205

Scoring table: BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Igapop 6.0 , Igapext 7.0

Searched: 2054640 seqs, 1451402878 residues

Total number of hits satisfying chosen parameters: 4107280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL tramer.p2n.model -DEV=slh
-Q=us-09-727-770-2 -L=1000000000 -L=1000000000 -L=1000000000
-DB=GenBank -DEMT=FASTA -SUFFIX=FASTA -MINMATCH=0.1 -FORMAT=FASTA
-UNITS=bits -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA
-LOCALIZATION -DB=FASTA -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA
-COFFI=FASTA -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA
-USER US0972770 -GCN 1.31616 -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA
-NO_XIPXY -NO_WMAP -LARGEST -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA
-WARR=1000000000 -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA -SUFFIX=FASTA
-YGAPE=10 -YGAPEXT=0.5 -YGAPEXT=0.5 -YGAPEXT=0.5 -YGAPEXT=0.5

Database: GenBank

- 1: qb.ba.*
- 2: qb.ba.*
- 3: qb.ba.*
- 4: qb.ba.*
- 5: qb.ov.*
- 6: qb.pat.*
- 7: qb.ph.*
- 8: qb.pl.*
- 9: qb.pr.*
- 10: qb.ro.*
- 11: qb.sts.*
- 12: qb.sy.*
- 13: qb.un.*
- 14: qb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mus.*
- 20: em.ov.*
- 21: em.ro.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*

- 29: em.vi.*
- 30: em.htq.hum.*
- 31: em.htq.inv.*
- 32: em.htq.other.*
- 33: em.htq.mus.*
- 34: em.htq.pln.*
- 35: em.htq.rood.*
- 36: em.htq.mam.*
- 37: em.htq.vrl.*
- 38: em.sr.*
- 39: em.htqo.hum.*
- 40: em.htqo.mus.*
- 41: em.htqo.other.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1056	100.0	643	6	AX337957 Sequence
2	1056	100.0	643	6	AX337959 Sequence
3	1056	100.0	151999	2	AF005474 Homo sapi
4	1056	100.0	104595	2	AF015914 Homo sapi
5	1043	98.8	139384	2	AC109333 Homo sapi
6	673	63.7	1078	9	BC007759 Homo sapi
7	673	63.7	1086	9	BC004537 Homo sapi
8	673	63.7	1123	9	BC009290 Homo sapi
9	673	63.7	1143	9	BC007389 Homo sapi
10	626.5	59.3	1162	9	BC004507 Homo sapi
11	616	58.3	1175	9	BC010147 Homo sapi
12	604	57.3	1153	10	MUSMUS
13	592.5	56.1	1133	10	MUSMUS
14	594.5	55.4	1092	10	MUSMUS
15	583	54.2	164263	2	AF094525 Homo sapi
16	569.5	53.1	216534	2	AF087122 Mus muscu
17	529	50.1	10813	10	AB055562 Mus muscu
18	529	50.1	186276	2	AF116716 Mus muscu
19	524.5	49.7	110903	9	BC0192817 Human DNA
20	516	49.3	114529	9	AF098526 Rattus no
21	488.5	46.3	146571	2	AF138376 Rattus no
22	458	43.4	468	4	AF027705 Ovis arie
23	451.5	42.8	6186	10	AF156088 Mus muscu
24	448	42.4	616	5	TM15KDM
25	436.5	41.3	210105	2	AF125948 Rattus no
26	427	40.4	902	5	AF099523 Bario rei
27	410.5	38.9	854	3	AF008924 Aedes aeg
28	398.5	37.7	1148	3	DMVHAP
29	393	37.2	1015	3	HELAIPASEV
30	392	37.1	1428	3	HEVAIP125
31	391.5	36.4	519	2	AF127150
32	381.5	36.1	2970	2	AF019221 Drosophila
33	381.5	36.1	8340	3	AF004438 Drosophila
34	381.5	36.1	15928	3	AF010565 Drosophila
35	381.5	36.1	18853	3	AF010043 Drosophila
36	381.5	36.1	18853	3	AF010043 Drosophila
37	381.5	36.1	281869	2	AF009544 Drosophila
38	381.5	36.0	918	3	AF009544 Drosophila
39	362	34.3	696	8	CNS0101CN
40	361.5	34.2	1385	8	CT002977
41	357.5	33.9	636	8	CNS0101CN
42	351.5	33.3	1007	8	SCHPASE
43	351.5	33.3	62643	8	SCHPASE
44	350	33.1	7252	3	DMCS90C
45	350	33.1	41909	2	AF018015 Drosophila

ALIGNMENTS

RESULT 1

QY 201 HisAlaValAspIle 205
 Dbb 610 CATGCTGTGCATC 624

RESULT 2
 AX337959
 LOCUS AX337959 Sequence 3 from Patent WO0194480.
 ACCESSION AX337959
 VERSION AX337959.1 GI:10128675
 KEYWORDS human,
 SOURCE Locality/Organisms
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.

REFERENCE
 AUTHORS Li, Z., Wei, M.B., Ketchum, K., and Beasley, R.M.
 TITLE Isolated human transporter proteins, nucleic acid molecules encoding human transporter proteins, and uses thereof
 JOURNAL Patent: WO 0194480-A 33-063-2001.
 APPLERA Corporation Robert A. Millman Assistant Secretary (US)

FEATURES
 source 1..6339
 Location/Organisms
 /organism "Homo sapiens"
 /dbx_ref "taxon:9606"

BASE COUNT	1613 a	1626 c	1609 g	1491 t
ORIGIN				

Alignment Scores:
 Pred. No.: 9,330-85 Length: 6,339
 Score: 1056.00 Matches: 20%
 Percent Similarity: 100.00% Conserved: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

03-03-727-773 2 { 205 } X AX337959 (1-6339)
 QY 1 Met.Ser.Asn.Ser.Pro.Glu.Tyr.Val.Leu.Val.PheThr.Ile.Ser.Cys.Val.Met.AluThr 20
 Db 3060 ATGGCCAAACAGCACGCCGAGATAGCGTTTGTTTTCATCTCACTCAAGCTCATATGCTTC 3059
 QY 21 Met.Val.Ser.Ser.Gly.Leu.Gly.Val.Ala.Cys.Ser.Y.Met.Ala.Lys.Asn.Gly.His.IleMet 40
 Db 3060 AAGGCTCCAGTCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4119
 QY 41 AlaMet.Ser.Val.Met.Trp.Proc.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu 60
 Db 3120 GCCATGCTGTCTCATGTCGCGCAGAGTAGATGATGATGATGATGATGATGATGATGATGATGATG 4179
 QY 61 Alagly.Ile.IleThr.Ile.Tyrltyr.Gly.Leu.Val.AlaAvalProthr.Val.AsnSer.Leu.Asn 80
 Db 4180 GTGTGTATCATCACCATCATGAGCTATGAGCTATGAGCTATGAGCTATGAGCTATGAGCTATGAG 4239
 QY 81 Asp.Asp.Asn.Ser.Leu.Tyrltyr.Ser.Ser.Phe.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu 100
 Db 3240 GATGACAACAGTGTCTATACAGCTTTCTTCACAGCTGGAGCTGAGCTGAGCTGAGCTGAGCTGA 4299
 QY 101 Alagly.Phe.Ala.Ile.Val.Ile.Val.Gly.Asp.Fthr.Gly.Ser.GlyThr.Ala.Ileu.Ileu.Ileu 120
 Db 3390 GGCGGGLTTTCGATGCTCATACAGCTTCGCTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGA 3359
 QY 121 Arg.Leu.Phe.Val.Gly.Met.Ile.Leu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu.Ileu 140
 Db 3460 CGACTATTGTATGACATGATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3419
 QY 141 Lys.Leu.Pro.Leu.Ser.Lys.Pro Thr.Ser.His.Sarg.Ile.Arg.Cys.Lys.Asp.Ile.Pro.Ser.Ser 160
 Db 3420 AACAGAGCCCTCTCAAACCCACACAGATCACAGAAATACGATGATAAACACACCTCTCTCTCA 3479
 QY 161 Phe.Arg.Asn.Gly.Ile.Pro.Val.His.Val.Val.Gly.Ser.Trp.Pro.Ser.Val.Val.Asp.Leu 180
 Db 3480 TTCGGGAACAACAGCTTCACACCATATGCTTGGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTCA 4549

REFERENCE AUTHORS TITLE JOURNAL	1 (bases 1 to 184595) Hirten, R., Nisbaum, C. and Landor, E. Human sapiens chromosome 17, clone RP11-11A22 unpublished	2 (bases 1 to 184595) Hirten, R., Nisbaum, C., Landor, E., Allen, N., Anderson, M., Baldwin, J., Barina, N., Becker, J. R., Beardsley, R., Collins, S., Coleman, M., Collins, S., Collymore, A., Brown, A., Castelle, A., Colaninno, M., Domingue, M., Donnellan, E., Doyle, M., Cooke, P., DeRellanes, K., Dwyer, K., Fink, K., Gade, D., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, K., Gade, D., Galan, J., Gardy, S., Grant, G., Hanes, R., Harford, A., Horton, L., Holland, J., Johnson, E., Jones, C., Kamm, J., Karatas, A., Klein, J., Lewicki, J., Liu, Q., Locke, K., MacDonald, P., Marquis, N., McEwan, P., McGuck, A., McKernan, K., Melanelli, J., McLennan, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, J., O'Donnell, J., O'Donnell, J., O'Donnell, J., Peterson, K., Pollard, V., Riley, R., Roy, A., Santos, K., Seeger, P., Stange, T., Thomas, N., Thompson, N., Subramanian, A., Talamas, J., Testa, S., Tirrell, A., Vassilov, H., Vo, A., Wheeler, J., Wu, X., Wymann, D., Ye, W., Zimm, A., and Zody, M.
REFERENCE AUTHORS TITLE JOURNAL	3 (bases 1 to 184595) Hirten, R., Nisbaum, C., Landor, E., Allen, N., Anderson, S., Barina, N., Beardsley, R., Becker, J., Beardsley, R., Camarata, J., Chang, J., Chazotte, R., Chopel, Y., Collymore, A., Cooke, A., Cooke, P., DeRellanes, K., Dwyer, K., Gade, D., Galan, J., Gardy, S., Grant, G., Hanes, R., Harford, A., Horton, L., Johnson, E., Jones, C., Kamm, J., Klein, J., Lewicki, J., Liu, Q., Locke, K., MacDonald, P., Marquis, N., McEwan, P., McGuck, A., McKernan, K., Melanelli, J., McLennan, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, J., O'Donnell, J., O'Donnell, J., O'Donnell, J., Peterson, K., Pollard, V., Riley, R., Roy, A., Santos, K., Seeger, P., Stange, T., Thomas, N., Thompson, N., Subramanian, A., Talamas, J., Testa, S., Tirrell, A., Vassilov, H., Vo, A., Wheeler, J., Wu, X., Wymann, D., Ye, W., Zimm, A., and Zody, M.	Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

Cooke, P., Le-Arellano, K., Iwar, K., Domingo, M., Domestiano, E., Reyes, C.,
Garcera, P., Fitzhugh, W., Forrest, J., Funke, R., Gupta, D.,
Galatola, J., Gardyna, S., Grant, G., Hatos, B., Hefford, A., Horton, L.,
Howland, J., Johnson, S., Jones, P., Kann, T., Karakas, A., Klein, L.,
Lacoste, J., Leckie, J., Lopez, K., Macdonald, P., Marquis, N.,
McEwan, P., McKay, K., McKenna, K., Melancon, J., Metcalfe, J.,
Morrow, J., Naylor, J., Norman, H., O'Connor, J., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Seelye, P.,
Sizemore, N., Strommen, N., Subramanian, A., Takamas, J.,
Telesayo, S., Tirrell, A., Vassiliou, H., Vu, A., Wheeler, J., Wu, X.,
Wyman, D., Ye-Wen, J., Zaman, A., and Zody, M.

Direct
Download
Submitted (17-Nov-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 184595)

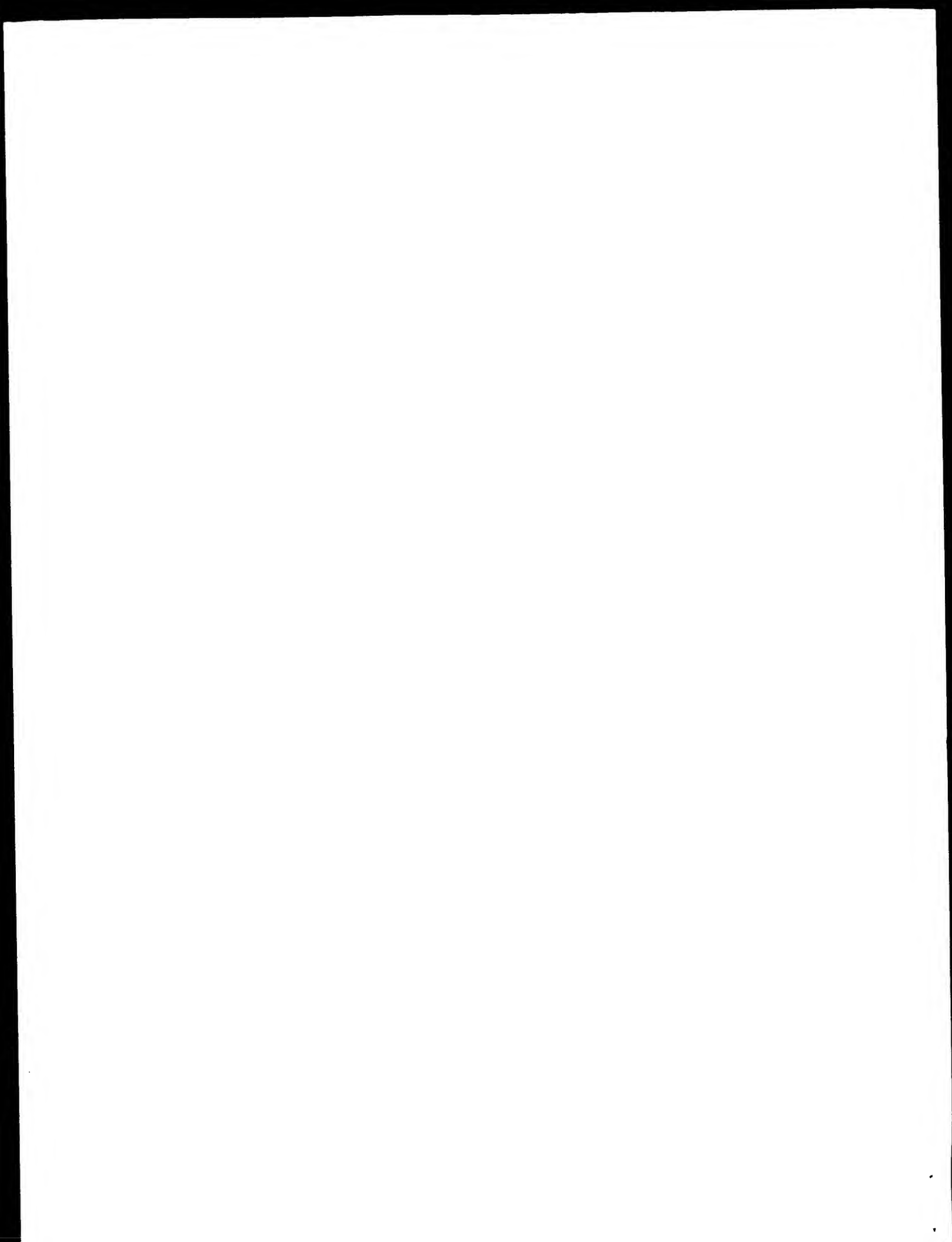
Hirton, B., Nusbaum, C., Landolt, E., Ali, A., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Blomster, P., Chapuis, Y., Collinmore, A.,
Camarato, J., Chang, J., Chazaro, B., Cooper, J., Dwyer, J., Edge, S.,
Cook, A., Cooke, P., De-Alcalá, K., Dewar, D., Galatola, J.,
Gard, S., Gerardi, P., Fitzgerald, M., Guo, D., Gutierrez, P.,
Faryna, S., Gord, S., Graham, R., Grant, P., Hart, N., Hayes, P.,
Horton, J., Hume, W., Hynes, J., Johnson, R., Jones, P., Kamel, A.,
Karakas, A., Kelly, C., Landolt, E., Levine, R., Lindblad-Toh, K.,

TITLE
JOURNAL

REFERENCE
AUTHORS

[illegible]

File time: 2657 secs



;	OTHER INFORMATION:	EXPRESSED IN ADULT LIVER.	4.4
;	OTHER INFORMATION:	EXPRESSED IN BRAIN.	15
;	OTHER INFORMATION:	EXPRESSED IN HEART.	5.1
;	OTHER INFORMATION:	EXPRESSED IN HBL100.	6.4
;	OTHER INFORMATION:		

```

Query Match      59.28; Score 375; Length 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 1.8e-102;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

y
368 CCGACTATTCTAG 382

b
434 CCGACTATTCTAG 448

```

RESULT 5
Sequence 49, Application US/09/25,401
Patent No. US2002095309A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, proteins and Antibodies
FILE REFERENCE: P106
CURRENT APPLICATION NUMBER: US/09/25,401
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US03/57862
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 50/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
1 SEQ ID NO 49
2 LENGTH: 956
3 TYPE: DNA
4 ORGANISM: Homo sapiens
5 FEATURE:
6 NAME/KEY: misc_feature
7 LOCATION: (37)
8 OTHER INFORMATION: n equals a t q. or c
9 NAME/KEY: misc_feature
10 LOCATION: (352)
11 OTHER INFORMATION: n equals a t q. or c
12 58-09-925-401-49

```

Query Match 45.3%; Score 287; DB 10; Length 956;
Best Local Similarity 77.4%; Pred. No. 4.7e-76;
Matches 401; Conservative 2; Mismatches 71; Indels 42; Gaps 3;
QY 12 GTCCACACACGCCGAGATACCTTTGGTTTCACATCTGGGTGATGGCCACCAT 71

[illegible]

```

RESULT 6
US-09-887-576-775
? Sequence 775, Application US/09887576
? Patent No. US20020144047A1
? GENERAL INFORMATION:
? APPLICANT: Budworth, P.
? APPLICANT: Budworth, P.
? APPLICANT: Chang, H.
? APPLICANT: Zhu, T.
? APPLICANT: Han, H.
? APPLICANT: Ward, X.
? APPLICANT: Cooper, Bret
? FILE OF INVENTOR: Promoters for regulation of plant expression
? FILE REFERENCE: 1360.001US1
? CURRENT APPLICATION NUMBER: US-09-887-576
? CORRECT FILING DATE: 2001-06-25
? PRIOR APPLICATION NUMBER: US-60/214,848
? PRIOR FILING DATE: 2000-06-24
? PRIOR APPLICATION NUMBER: US-60/214,887
? PRIOR FILING DATE: 2000-06-24
? PRIOR APPLICATION NUMBER: US-60/258,692
? PRIOR FILING DATE: 2000-12-29
? NUMBER OF SEQ ID NOS: 875
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 775
? LENGTH: 500
? TYPE: DNA
? ORGANISM: Oryza sativa
US-09-887-576-775

```

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Computer I.D.

M protein nucleic search, using frame_plus_p2n model

Run on: December 28, 2002, 10:56:01 : Search time 1974 seconds
(without alignments)
1681.902 Million cell updates/sec

File: US-09-727-770-2
Perfect score: 1056
Sequence: 1 HSNRPYALVFIIISAMAT.....PENVRLARVLIIFHVDI 205

Scoring Table:
BL0SUM62
Xgap 10.0 Xopen 0.6
Ygap 10.0 Yopen 0.5
Fgap 6.0 Fopen 7.0
Delop 6.0 Delopt 7.0

Searched: 1615466 seqs, 80774379 residues

Total number of hits satisfying chosen parameters: 3266032

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n_model -DB=xbh
-O=us09-727-770-2 -S=1 -L=1000000000 -M=1000000000 -N=1000000000 -P=1000000000 -Q=1000000000 -R=1000000000 -T=1000000000 -U=1000000000 -V=1000000000 -W=1000000000 -X=1000000000 -Y=1000000000 -Z=1000000000
-DB=EST -QFMT=lastal -SUFFIX=p2n -T=1 -MINMATCH=61 -FMT=1 -L=1000000000 -M=1000000000 -N=1000000000 -P=1000000000 -Q=1000000000 -R=1000000000 -T=1000000000 -U=1000000000 -V=1000000000 -W=1000000000 -X=1000000000 -Y=1000000000 -Z=1000000000
-UNITS=bits -STAPL=1 -FMT=1 -MATPIX=bl0sum62 -SPANS human0.041 -LIST=45
-LOCALIGN=200 -THE_SCOPE=ptl -THE_MAX=100 -THE_MIN=0 -ALIGN=15 -MCODE=LOCAL
-OUTFMT=ptl -NRM=ext -HPAFS=500 -MTNPN=0 -MAXIPN=2000000000
-USERS 500000000 -GCR 1.1856 -FMT=1 -L=1000000000 -S=1 -M=1000000000 -N=1000000000 -P=1000000000 -Q=1000000000 -R=1000000000 -T=1000000000 -U=1000000000 -V=1000000000 -W=1000000000 -X=1000000000 -Y=1000000000 -Z=1000000000
-NO_XIPV -NO_WMAP -LAPCHQUERY -NET_SCORES=0 -WAL=1 -L=1000000000 -DEV_LIMB=120
-WARN_LIMB=120 -THREADS=1 -XAPR=10 -XAPRPT=0 -XAPRPT=0 -XAPRPT=0 -XAPRPT=0
-YAPR=10 -YAPRPT=0 -DELIF 5 -DELIF 5 -DELIF 5

Database :
EST :
1: em_estha1*
2: em_estha2*
3: em_estha3*
4: em_estha4*
5: em_estha5*
6: em_estha6*
7: em_estha7*
8: em_estha8*
9: em_estha9*
10: em_estha10*
11: em_estha11*
12: em_estha12*
13: em_estha13*
14: em_estha14*
15: em_estha15*
16: em_estha16*
17: em_estha17*
18: em_estha18*
19: em_estha19*
20: em_estha20*
21: em_estha21*
22: em_estha22*
23: em_estha23*
24: em_estha24*
25: em_estha25*
26: em_estha26*
27: em_estha27*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	674.5	61.9	1061	14	BM562656	BM562656 AGNCOURT
2	673	61.7	908	14	BQ722145	BQ722145 AGNCOURT
3	673	61.7	981	14	BQ678034	BQ678034 AGNCOURT
4	673	61.7	994	14	BM463421	BM463421 AGNCOURT
5	673	61.7	1063	14	BQ353546	BQ353546 AGNCOURT
6	673	61.7	1081	14	BQ232987	BQ232987 AGNCOURT
7	673	61.7	996	14	BQ068594	BQ068594 AGNCOURT
8	666	61.1	986	14	BQ232102	BQ232102 AGNCOURT
9	666	61.1	923	14	BQ424541	BQ424541 AGNCOURT
10	664	62.9	1076	14	BQ229405	BQ229405 AGNCOURT
11	664	62.9	1201	14	BM473665	BM473665 AGNCOURT
12	662	62.7	991	14	BM918349	BM918349 AGNCOURT
13	660	62.5	1031	14	BQ273185	BQ273185 AGNCOURT
14	660	62.5	1031	14	BM554691	BM554691 AGNCOURT
15	659	62.4	1076	14	BM544598	BM544598 AGNCOURT
16	659	62.4	1222	14	BM478359	BM478359 AGNCOURT
17	654	61.9	872	14	BQ722338	BQ722338 AGNCOURT
18	654	61.8	896	14	BM791194	BM791194 K-EST0071
19	651	61.6	908	14	BQ233921	BQ233921 AGNCOURT
20	648	61.4	631	14	BM769410	BM769410 K-EST0052
21	648	61.4	634	14	BM843350	BM843350 K-EST0121
22	648	61.4	638	14	BM818948	BM818948 K-EST0086
23	648	61.4	657	14	BM769872	BM769872 K-EST0053
24	644	61.3	936	14	BQ919999	BQ919999 AGNCOURT
25	644	61.3	1093	14	BM547917	BM547917 AGNCOURT
26	641	60.7	983	14	BQ722145	BQ722145 AGNCOURT
27	638.5	60.5	1017	14	BM549877	BM549877 AGNCOURT
28	637	60.3	907	14	BQ878221	BQ878221 AGNCOURT
29	636.5	60.3	913	14	BQ949611	BQ949611 AGNCOURT
30	635	60.2	995	14	BM921084	BM921084 AGNCOURT
31	635	60.2	1119	13	BM450080	BM450080 AGNCOURT
32	634	60.0	991	9	AL524214	AL524214
33	634	60.0	940	14	BQ719555	BQ719555 AGNCOURT
34	628	59.5	969	14	BQ060051	BQ060051 AGNCOURT
35	627	59.4	1169	14	BM928304	BM928304 AGNCOURT
36	624.5	59.3	1136	13	BM554312	BM554312 AGNCOURT
37	623.5	59.0	904	14	BQ218416	BQ218416 AGNCOURT
38	622.5	58.9	1090	13	BM453233	BM453233 AGNCOURT
39	620	58.7	641	12	BQ297409	BQ297409 AGNCOURT
40	620	58.7	670	12	BQ741416	BQ741416 AGNCOURT
41	620	58.7	722	12	BQ797916	BQ797916 AGNCOURT
42	618	58.5	1104	14	BM924179	BM924179 AGNCOURT
43	617.5	58.5	714	10	BM251775	BM251775 AGNCOURT
44	615	58.2	654	14	BM948591	BM948591 U1-M-PMO
45	612	58.0	682	14	BQ831675	BQ831675 AGNCOURT

ALIGNMENTS

RESULT 1
BM562656
LOCUS BM562656
DEFINITION AGNCOURT_6566751 NIH_MGC_H8 Homo sapiens cDNA clone IMAGE:573684
5', mRNA sequence.
ACCESSION BM562656
VERSION BM562656.1 GI:18808940
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1061)
AUTHORS NIH-MGC http://www.ncbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

Tue Dec 31 14:03:17 2002

```

VERSION RM464421.1 GI:18512463
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 993)
AUTHORS NIH MGC http://www.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstrauss@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM2352 row: k column: 23
High quality sequence stamp: 711.
Location/Qualifiers
1..993
/orqanism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5585806"
/clone_lib="NIH_MGC_92"
/clone_type="embryonic"
/issue="type:"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site: 1; Note:
Site_2: Salt: Cloned and direct sequenced; oligo-dT primed
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 159 a 369 c 259 g 206 t
ORIGIN
Alignment Scores:
Pred No.: Length: 993
Score: 673.00 Matches: 152
Percent Similarity: 74.19% Conservatives: 9
Best Local Similarity: 70.05% Mismatches: 42
Query Match: 63.73% Indels: 15
DB: 13 Gaps: 3
US-09-727-770-2 (1-205) x RM464421 (1-993)
QY 2 SerAsuSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThrMet 21
Db 101 TCGAAGAGGGGGCGGAGTATGCTTCTTTTCGGGTCATGGGGGCGGCGGCGGCGG 160
QY 22 ValSerSerGlyPheGlyAlaAlaGlySerGlyPheAlaMetAlaMetAlaAla 41
Db 161 GTCTTCAGGCGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 220
QY 42 MetSerValMetTrpProGluLeuIleLeuMetLeuSerIleIleProValValMetAla 61
Db 221 ATGCTGCTCATGCGGGCGGAGAGAGC---ATGAAGTCCATCATCATCATCATCAT 277
QY 62 GlyTleIleThrIleTyrGlyLeuValAlaAlaValProProAlaAsuSerLeuAsp 81
Db 278 GGCATCATGCGCTCATGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 337
QY 82 AspAsuSerLeuTyrSerSerPheLeuGluLeuGlyAlaValAlaValAlaValAla 97
Db 438 GACATACAGCGCTCTACAGAGAGTCTGCTGAGCTGAGGCGGCGGCGGCGGCGGCG 397
QY 98 GlyLeuAlaGlyPheAlaIleValIleValGlyAspThrGlyGlySerGlyThrAla 117
Db 498 GCGCTGGGAGCGCGCTTGGATCATGGGAGTGGTGGGAGCGTGGGCGGCGGCGGCG 457
QY 118 GluGlnProAcuLeuPheValGlyMetIleIleIleLeuIlePheAlaLysVal----- 135

```

```

458 CAGCAGCGCGGACATATTCGTCGGGATGATGATCTGATCTCATCTTTCGACGCTCTGAG 517
QY 146 -----LeuIleLeuSerThrIleSerThrIleSerThrIleSerThrIle 149
Db 518 CTCACAGGCTGCTACGTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577
QY 149 SerHisArgIleAlaCysLysAspHisProSerGlyAlaAsuLysIleIleAspThr 168
Db 578 AGCCACAGAAATATTATTAACATACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 637
QY 169 HisValLeuGlySerTrpProSerValValAspLeuLeuSerValGlyGlySerAla 188
Db 638 TACGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 696
QY 189 HisArgLeuLeuAlaArgProCysProLeuLeuProIleHisAlaValAspIle 205
Db 697 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747
RESULT 5
QY05:546
LOCUS AGNC356KT.6845437 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:554542
5', mRNA sequence.
ACCESSION Q0053546
VERSION Q0053546.1 GI:19812886
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1063)
AUTHORS NIH MGC http://www.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstrauss@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DRS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM2122 row: o column: 23
High quality sequence stamp: 692.
Location/Qualifiers
1..1063
/orqanism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:554542"
/clone_lib="NIH_MGC_106"
/clone_type="natural"
/notes="Organ: natural killer cells; cell line"
Note: cDNA made by oligo dT priming, directionally cloned
into EcoRI/XbaI sites using the following 5' adaptor:
GACGACGAC(G). Library constructed by Lind Bond in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using zap cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 176 a 370 c 401 g 215 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: Length: 1063
Score: 673.00 Matches: 152
Percent Similarity: 74.19% Conservatives: 9
Best Local Similarity: 70.05% Mismatches: 42
Query Match: 63.73% Indels: 15
DB: 14 Gaps: 3
US_09_727_770_2 (1-205) x Q0053546 (1-1063)

```

```

QY 2 SetAsnAsuSerProAlaLeuValPheThrIleSerGlyAlaMetAlaThrMet 21
DB 85 TCAAGAGAGCCGCCAGTAAGTTCGTTTCGGGTCATCGGCGGCTCGGCGGCATG 144
QY 27 ValSerSerGlyLeuGlyAlaIleGlySerGlyMetAlaIleGlySerGlyIleMetAla 41
DB 145 GTCCTAGAGCCGCTAGGCGCTAGGTAAGGACAGCAAGCAAGCAAGCAAGCAAGCAAG 204
QY 42 MetSerValIleThrIlePheThrIleGlyMetAlaIleGlySerGlyIleMetAla 61
DB 205 ATGCTGTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 261
QY 62 GlyIleThrIleThrIleGlyLeuValAlaValProAlaAsuSerLeuAsuAsp 81
DB 262 GCAATCATCCCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
QY 82 AspAsuSerLeuThrIlePheThrIleGlyMetAlaIleGlySerGlyIleMetAla 101
DB 322 GATATCAGGCTCTACAAAGAGCTTCTCCAGCTGCGCGCGCGCGCGCGCGCGCGCG 481
QY 98 GlyLeuAlaValAcylPheAlaIleValIleValIleValIleValIleValIleVal 117
DB 382 GGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 441
QY 118 GlyIlePheArgLeuPheValGlyMetIleLeuIleLeuIlePheAlaIleGlyVal 135
DB 442 CACACGCCCGGACATCGTGGCGCAAGATGATGATGATGATGATGATGATGATGATG 501
QY 146 -----LeuIleLeuSerThrLysGlnProLeuSerThrProThr 148
DB 502 CTGTAGCGGTCTCATCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 561
QY 149 SerHisAsuIleAlaGlySerGlySerGlySerGlySerGlySerGlySerGlySer 168
DB 512 AGACACAGAAATATATGAAAGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
QY 169 HisValLeuGlySerThrPheSerValIleAspLeuSerValIleGlySerThrVal 188
DB 622 TACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 680
QY 189 HisArgLeuAlaValArgProGlySerProLeuProGlyHisAlaValAspIle 205
DB 681 CATAGCTGTTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 741

RESULTS
LOCUS RQ242887 1081 bp mRNA linear EST 02 MAY 2002
DEFINITION AGENCOURT_7573952 NIH_MRG_99 Homo sapiens cDNA clone IMAGE:5814468
5' mRNA sequence.
ACCESSION RQ242887
VERSION RQ242887.1 GI:20414287
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
XREF GenBank: RQ242887
COMMENT
  cDNA library prepared by: Life Technologies, Inc.
  cDNA library arrayed by: The I.M.A.G.E. Consortium (LIND)
  DNA sequenced by: Amersham Bioscience Corporation
  clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/BLI at:
  http://image.llnwd.net
  plate: L14M13423 row: d column: 12
  High quality sequence stop: 792.
  Location/Qualifiers

```

```

Source 1. 1081
/organism="Homo sapiens"
/job_xref="taxon:9606"
/clone_lib="IMAGE:6058431"
/clone_id="NIH_MGC_68"
/analysis_type="large scale sequencing"
/lab_host="BIOIB (phase-resistant)"
/vector="puc19, long vector, pCMV-Sp6K16, Site 1, NotI"
Site_2: Salt; Cloned unidirectionally. primer: oligo dT.
Average insert size: 1.4 kb library constructed by Life
Technologies."
BASE COUNT 185 a 383 c 285 g 227 t 1 others
ORIGIN
Alignment Scores: 1,83e-63 Length: 1081
Prod. No.: 673,00 Matches: 152
Score: 74.19% Conservativity: 9
Best Local Similarity: 70.05% Mismatches: 42
Query Match: 63.73% Indels: 15
DB: 14 Gaps: 3
US-09-727-770-2 (1-205) x HQ242887 (1-1081)
QY 2 SetAsnAsuSerProGlyThrAlaLeuValPheThrIleSerGlyAlaMetAlaThrMet 21
DB 83 TCAAGAGAGCCGCCAGTAAGTTCGTTTCGGGTCATCGGCGGCTCGGCGGCATG 142
QY 27 ValSerSerGlyLeuGlyAlaIleGlySerGlyMetAlaIleGlySerGlyIleMetAla 41
DB 143 GTCCTAGAGCCGCTAGGCGCTAGGTAAGGACAGCAAGCAAGCAAGCAAGCAAGCAAG 202
QY 42 MetSerValIleThrIlePheThrIleGlyMetAlaIleGlySerGlyIleMetAla 61
DB 203 ATGCTGTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 259
QY 62 GlyIleThrIleThrIleGlyLeuValAlaValProAlaAsuSerLeuAsuAsp 81
DB 260 GCAATCATCCCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319
QY 82 AspAsuSerLeuThrIlePheThrIleGlyMetAlaIleGlySerGlyIleMetAla 97
DB 320 GATATCAGGCTCTACAAAGAGCTTCTCCAGCTGCGCGCGCGCGCGCGCGCGCGCG 379
QY 98 GlyLeuAlaValAcylPheAlaIleValIleValIleValIleValIleValIleVal 117
DB 380 GGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 439
QY 118 GlyIlePheArgLeuPheValGlyMetIleLeuIleLeuIlePheAlaIleGlyVal 135
DB 440 CACACGCCCGGACATCGTGGCGCAAGATGATGATGATGATGATGATGATGATGATG 499
QY 136 -----LeuIleLeuSerThrLysGlnProLeuSerThrProThr 148
DB 500 CTGTAGCGGTCTCATCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 559
QY 149 SerHisAsuIleAlaGlySerGlySerGlySerGlySerGlySerGlySerGlySer 168
DB 512 AGACACAGAAATATATGAAAGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619
QY 169 HisValLeuGlySerThrPheSerValIleAspLeuSerValIleGlySerThrVal 188
DB 620 TACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 678
QY 189 HisArgLeuAlaValArgProGlySerProLeuProGlyHisAlaValAspIle 205
DB 679 CATAGCTGTTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 729
RESULTS
LOCUS RQ058594 996 bp mRNA linear EST 24-MAR-2002
DEFINITION AGENCOURT_6824891 NIH_MRG_99 Homo sapiens cDNA clone IMAGE:5814468
5' mRNA sequence.

```

[illegible]

Tue Dec 31 14:03:17 2002

```

5' mRNA sequence.
BQ229305
VERSION 1 GI:20410705
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1076)
RefSeq Project: 1.3.4
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rstrauss@mail.nih.gov
Tissue Procurement: pcpm/ghp/azdar
cDNA library preparation: Life Technologies, Inc.
cDNA library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAM1323 row: a column: 15
High quality sequence stop: 715.
FEATURES
source
1..1076
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="BHL10B (phage-resistant)"
/notes="Organ: Lung; Vector: pCMV-Sport6; Site: 1: NCI;
Site: 2: Salt; Cloned unidirectionally. Primer: Oligo dt.
Average insert size: 1.8 kb. Library constructed by Life
Technologies. Note: this is a NIH MGC library."
BASE COUNT 169 a 391 c 297 g 218 t 1 others
ORIGIN
1
1076
Alignment Scores:
Pred. No.: 1,740-62 Length: 1076
Score: 664.00 Matches: 151
Percent Similarity: 73.73% Conservat: 9
Best Local Similarity: 69.59% Mismatches: 43
Query Match: 62.88% Indels: 15
Gaps: 3
US 09-727-770-2 (1-205) x BQ229305 (1-1076)
QY 2 SerAsnAsnSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThrMet 21
DB 122 TCCAGACGGCGCGAGTATGCTTCGTTTTCGGCGGTCATGGCGGCTCGCGCGGATG 181
QY 22 ValSerSerGlyLeuGlyAlaLeuValPheThrIleSerGlyMetAlaLeuValPheThrIleSerGly 41
DB 182 GTCTTCACAGCGCTGGCGGCTGCTATGCAACAGCAGAGAGAGTACGGGATTCGGCGCC 241
QY 42 MetSerValMetThrTrpProGluMetIleHisSerIleLeuPheValValMetAla 61
DB 242 ATGTCGTGTCATGGCGCGGCGAGATGTC---AUGAAGTCATGATGCCAGTGGTCAAGGCT 298
QY 62 GlyIleIleThrIleTyrGlyLeuValAlaLeuValProProValAsnSerLeuAsp 81
DB 299 GACATCATGGGATCATGGCTGGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358
QY 82 AspAsnSerLeuTyrSerSerPheLeuGluGluGlyAlaLeuValPheThrIleSerGly 97
DB 359 GACATCATGGGATCATGGCTGGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418
QY 98 GlyLeuAlaLeuIlePheAlaIleValIleValIlePheThrIlePheGlyThrAla 117
DB 419 GCGCTGGAGGCGGCTTCCCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478
QY 118 GluGlnProArgLeuPheValIleIlePheIleLeuValPheAlaLeuVal----- 125

```

```

|||||
479 CAGCAGCCGCACTATTCGTGGCAAGCAAGCTGATCTTCTTCGCGCAAGCAAGCAAGCAAG 548
QY 136 -----LeuIleLeuSerThrLysThrProLeuSerLysProThr 148
DB 539 CTCTACGGCTCTCATCGCGCGCTCATGCTCTCTCCCAAAAGTAGACCTTCCTCGAGCTCACT 598
QY 149 SerHisArgIleArgCysLysAspHisProSerSerPheAlaGAsulysLeuProAspHis 168
DB 599 AGCTACAGAAATATTATTAAGAGACACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 658
QY 169 HisValLeuGlySerTrpProSerValValAspLeuLeuSerValGluGlnCysProValVal 188
DB 659 TACGCAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 217
QY 189 HisArgLeuLeuAlaArgProCysProLeuProProHisAlaValAspIle 205
DB 718 CATGGTCTGTTCTCGCGCGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 768
RESULT 11
HMA73665 1261 bp mRNA linear EST 05 FEB 2002
ACCESSION AGCNC00127.6466485 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5562054
VERSION HMA73665
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
RefSeq Project: 1.3.4
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rstrauss@mail.nih.gov
Tissue Procurement: ATCC
cDNA library preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAM12290 row: a column: 06
High quality sequence stop: 690.
FEATURES
source
1..1201
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_98"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="BHL10B (phage resistant)"
/notes="Organ: small intestine; Vector: pCMV-Sport6;
Site: 1: Not 1; Site: 2: Salt; Cloned unidirectionally.
Oligo-dt primed. Average insert size 1.767 kb. Library
constructed for full length clones and constructed by Life
Technologies. Note: this is a NIH MGC library."
BASE COUNT 200 a 425 c 419 g 256 t 1 others
ORIGIN
1
1201
Alignment Scores:
Pred. No.: 2,140-62 Length: 1201
Score: 664.00 Matches: 148
Percent Similarity: 74.18% Conservat: 10
Best Local Similarity: 69.48% Mismatches: 40
Query Match: 62.88% Indels: 15
Gaps: 4
US 09-727-770-2 (1-205) x HMA73665 (1-1201)
QY 2 SerAsnAsnSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThrMet 21

```



```

DB: 13 gaps: 3
US-09-727-770-2 (1-205) x BM544491 (1-1041)
QY 2 SerAsuAsuSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThrMet 21
DB 94 TCAAGAGCGCCCGGAGTATATCTCTTTGGGCTTAATGGGCGCGCGGCGGCGATG 153
QY 22 ValSerSerGlyLeuGlyAlaAlaCysGlyMetAlaLysAsuGlyThrGlyMetAla 41
DB 154 GCTTCAGCGCCCTGGAGCTTCTATGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG 213
QY 42 MetSerValMetTrpProGlyGluIleHisMetLysSerIleLeuValMetAla 61
DB 214 ATGCTCTCAAGCGCGCGGAGGAGATG---ATGAAATGTAATGAGGAGGAGGAGGAG 279
QY 62 GlyIleIleThrIleTyrGlyLeuValAlaValProProAlaAsuSerLeuAsp 81
DB 271 GATCAICAGCGCATCTACGGCGCTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 330
QY 82 AspAsuSerLeuTyrSerSerPheLeuGluIleLeuValAla-----GlyLeuSer 97
DB 531 CATACAGGCTCTACAGAGCTCTCTGAGTGTGGAGGAGGAGGAGGAGGAGGAGGAG 390
QY 98 GlyLeuAlaAlaGlyPheAlaIleValIleValIleValAspThrGlyLysGlyThrAla 117
DB 591 GAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 459
QY 118 GluIleProAlaLeuPheValGlyMetIleLeuIleLeuIlePheAlaLysVal----- 135
DB 451 CAGCAGCGCGGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 510
QY 136 -----LeuIleLeuSerThrIleValProLeuSerLysProThr 148
DB 511 CTCTACGGCTCATCTGCGCGCTCATCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 570
QY 149 SerHisArgIleArgGlyLysAspHisProSerGlyPheAlaAsuGlyGlnProAspThr 169
DB 571 AGGACAGAATATATGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 630
QY 169 HisValLeuGlySerTrpProSerValValAspLeuSerValIleCysSerArgVal 189
DB 631 TACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 689
QY 189 HisArgLeuLeuAlaArgProGlySerProLeuProPheAlaValAspIle 205
DB 648 CATGCTGTGTTTCCCGCGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741

RESULTS
BM544491
LOCUS: 1076 bp mRNA 1076 bp EST 20 FEB 2002
DEFINITION: ACDH30PT_5492852 NM_00124 Homo sapiens cDNA clone IMAGE:5728258
5' mRNA sequence.
VERSION: BM544498
KEYWORDS: EST.
SOURCE: Human.
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1076)
GenBank accession: U01414
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausfeld, Ph.D.
Email: cyabbs@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

```

```

Plate: L1AM12724 row: k column: 11
High quality sequence start: 47
High quality sequence stop: 733.
location/Qualifiers
1..1076
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAG:5728258"
/contig="114"
/tissue_type="hippocampus"
/lab_host="DH10B"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site: 1: EcorV
(site 2: Koll. RNA source male hippocampus,
site 2: library is oligo-dT primed and directionally
cloned (EcorV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full length clones and was
constructed by G. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 182 a 393 c 275 g 225 t 1 others
ORIGIN
Alignment Scores: 6,460-62 Length: 1076
Pred. No.: 659,90 Matches: 149
Score: 73,849 Conservative: 9
Percent Similarity: 69,638 Mismatches: 42
Best Local Similarity: 62,418 Indels: 15
Query Match: 13 Gaps: 3
US-09-727-770-2 (1-205) x BM544498 (1-1076)
QY 2 SerAsuAsuSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThrMet 21
DB 145 TCAAGAGCGCCCGGAGTATATCTCTTTGGGCTTAATGGGCGCGCGGCGGCGATG 204
QY 22 ValSerSerGlyLeuGlyAlaAlaCysGlyMetAlaLysAsuGlyThrGlyMetAla 41
DB 205 GTTCTCAAGCGCATCTACGGCGCTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 264
QY 42 MetSerValMetTrpProGlyGluIleHisMetLysSerIleLeuValMetAla 61
DB 265 ATGCTCTCAAGCGCGGAGGAGATG---ATGAAATGTAATGAGGAGGAGGAGGAG 321
QY 62 GlyIleIleThrIleTyrGlyLeuValAlaValProProAlaAsuSerLeuAsp 81
DB 322 GATCAICAGCGCATCTACGGCGCTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 381
QY 82 AspAsuSerLeuTyrSerSerPheLeuGluIleLeuValAla-----GlyLeuSer 97
DB 382 GAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 441
QY 98 GlyLeuAlaAlaGlyPheAlaIleValIleValIleValAspThrGlyLysGlyThrAla 117
DB 442 GGCCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 501
QY 118 GluIleProAlaLeuPheValGlyMetIleLeuIleLeuIlePheAlaLysVal----- 135
DB 502 CAGCAGCGCGGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 561
QY 136 -----LeuIleLeuSerThrIleValProLeuSerLysProThr 148
DB 562 CTCTACGGCTCATCTGCGCGCTCATCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
QY 149 SerHisArgIleArgGlyLysAspHisProSerGlyPheAlaAsuGlyGlnProAspThr 168
DB 622 AGGACAGAATATATGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
QY 169 HisValLeuGlySerTrpProSerValValAspLeuSerValIleCysSerArgVal 188
DB 682 TACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 740
QY 189 HisArgLeuLeuAlaArgProGlySerProLeuProPheAlaValAspIle 202

```

Db 741 CATGTCGTTTCGGGGGGCTTACCGGGGGGGCGTGGC 782

Search completed: December 29, 2002, 12:19:27
Job time : 1978 secs

GenCore version 5.1.4
Copyright (c) 1994 - 2002 Computer Ltd.

EM protein - nucleic search, using frame_plus_p2n model

Run on: December 28, 2002 16:57:11 - search time 51 seconds
(without alignments)
1242229 9011 0 511 4-1450 5/50

Title: US-09-727-770-2
Perfect score: 1056
Sequence: 1 MSNNSPEVALYETISGMAT.....PEVHSLAPPCHLPHAVCI 205

Scoring table: nucscwm2

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441162 seqs, 153113391 residues

Total number of hits satisfying chosen parameters: 382724

Minimum hit seq length: 0

Maximum hit seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -PEV-xlh

-C 1242229 9011 0 511 4-1450 5/50

-DB-Issued_Patents_NA -QEMT=fastap -SUFFIX=p2n.rni -MINMATCH 0.1 -LOOPCL=0

-LO-PEXT-0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS human40.cdi

-LIST=45 -DBALIGN=200 -THE_SCORE=pe -THE_MAX=100 -THE_MIN=0 -ALIGN=15

-MQUF=LOCAL -OUTFMT=pt6 -NORM=ext -HEAUS17P=500 -MINLEN=0 -MAXLEN=200000000

-USER=050972779.355N 1.1.17 -result 27122002 394449 5870 -NOPS 6 -ICPU 3

-W 33345 9 3963 -A 1000000 1000000000 -WA 1000000000 -MAX_LEN=0 130

-WARN_TIMEOUT=40 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop 6 -Fgapext=7

-Ygapop=10 -Ygapext=0.5 -DELOP 6 -DELEXT 7

Database:

Issued_Patents_NA:

1: 1242229 9011 0 511 4-1450 5/50
2: 27122002 394449 5870 6 1000000000
3: 33345 9 3963 -A 1000000 1000000000 -WA 1000000000 -MAX_LEN=0 130
4: 1242229 9011 0 511 4-1450 5/50
5: 1242229 9011 0 511 4-1450 5/50
6: 1242229 9011 0 511 4-1450 5/50

Note: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	134.5	12.7	4982	US-08-699-1048-1	Sequence 1, Appl1
2	134.5	12.7	4982	US-09-229-059-1	Sequence 1, Appl1
3	134.5	12.6	849	US-08-807-300-2	Sequence 2, Appl1
4	100.5	9.5	697	US-08-254-494-2	Sequence 2, Appl1
5	100.5	9.5	687	US-08-408-222B-2	Sequence 2, Appl1
6	100.5	9.5	1120	US-08-408-222B-4	Sequence 3, Appl1
7	94.5	9.1	1120	US-08-254-494-4	Sequence 3, Appl1
8	89.5	8.5	3807	US-08-816-765-1	Sequence 1, Appl1
9	89.5	8.5	3807	US-08-090-673-1	Sequence 1, Appl1
10	84.5	8.0	31571	US-08-323-449B-1	Sequence 2, Appl1
11	84.5	8.0	53526	US-08-658-146-2	Sequence 2, Appl1
12	84.5	8.0	53527	US-08-658-146-1	Sequence 1, Appl1

C 13	84	8.0	21706	4	US-08-761-527-36	Sequence 46, Appl1
C 14	81.5	7.7	3524	3	US-08-801-444-7	Sequence 7, Appl1
C 15	91.5	7.7	9524	4	US-09-199-696-7	Sequence 7, Appl1
C 16	80	7.6	3077	1	US-08-703-809-1	Sequence 1, Appl1
C 17	80	7.6	3077	1	US-08-703-808-1	Sequence 1, Appl1
C 18	80	7.6	3077	2	US-08-914-066-1	Sequence 1, Appl1
C 19	80	7.6	3077	2	US-08-703-807-1	Sequence 1, Appl1
C 20	80	7.6	3077	2	US-08-703-807-1	Sequence 1, Appl1
C 21	80	7.6	3077	3	US-09-211-631-1	Sequence 1, Appl1
C 22	80	7.6	3077	4	US-05-555-528-1	Sequence 1, Appl1
C 23	80	7.6	3077	4	US-09-001-141-1	Sequence 1, Appl1
C 24	80	7.6	3077	4	US-05-532-803-1	Sequence 1, Appl1
C 25	80	7.6	3077	4	US-09-653-803-3	Sequence 3, Appl1
C 26	80	7.6	3077	4	US-10-013-784-3	Sequence 3, Appl1
C 27	79.5	7.5	10684	4	US-09-221-017B-401	Sequence 401, Appl1
C 28	79.5	7.5	410666	4	US-09-103-840A-2	Sequence 2, Appl1
C 29	79.5	7.5	411529	4	US-09-103-840A-1	Sequence 17, Appl1
C 30	79	7.5	491	4	US-09-189-060B-17	Sequence 17, Appl1
C 31	79	7.5	4739	4	US-08-677-970-1	Sequence 1, Appl1
C 32	79	7.5	7766	4	US-09-125-619-3	Sequence 3, Appl1
C 33	78.5	7.4	982	4	US-08-961-527-309	Sequence 309, Appl1
C 34	78.5	7.4	2153	2	US-08-801-801-2	Sequence 2, Appl1
C 35	77.5	7.3	4294	1	US-08-152-494B-2	Sequence 2, Appl1
C 36	77	7.3	2344	1	US-07-695-472B-35	Sequence 35, Appl1
C 37	77	7.3	2347	1	US-08-250-740-32	Sequence 32, Appl1
C 38	77	7.3	2347	1	US-07-695-472B-1	Sequence 1, Appl1
C 39	77	7.3	2347	1	US-07-695-472B-34	Sequence 34, Appl1
C 40	77	7.3	2347	1	US-07-695-472B-36	Sequence 36, Appl1
C 41	77	7.3	15239	1	US-08-390-879-17	Sequence 17, Appl1
C 42	76.5	7.2	2249	2	US-08-860-635A-18	Sequence 18, Appl1
C 43	76.5	7.2	2249	4	US-09-281-476-18	Sequence 18, Appl1
C 44	76	7.2	1254	4	US-09-030-267-4	Sequence 4, Appl1
C 45	76	7.2	1316	1	US-08-579-667-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1

US-08-699-103B-1

Sequence 1, Application US/08699103B

Patented by: 357422

GENERAL INFORMATION:

APPLICANT: Rinc, Jasper D.

APPLICANT: Hampton, Randolph

TITLE OF INVENTION: CIRCUITS AND METHODS FOR CONTROLLING

FUNCTION OF INVENTION: CIRCUITS FOR SYNTHESIS

INVENTOR: Rinc, Jasper D.

INVENTOR: Hampton, Randolph

ADDRESS: 4200 Sand Hill Road, Suite 100

City: Menlo Park

State: CA

Country: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/699, 103B

FILING DATE: 16-AUG-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/002,581

FILING DATE: 17-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: Green, Grant D.

REGISTRATION NUMBER: 31,259

REFERENCE/INVENTION NUMBER: 09/727,000/5001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/322,5070

TELEFAX: 650/854-0875

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 4982 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORGANISM: *Yersinia enterocolitica*

1. CORRESPONDENCE ADDRESS:
 2. ADDRESSEE: Dike, Bronstein, Roberts & Cashman
 3. STREET: 130 Water Street
 4. CITY: Boston
 5. STATE: MA
 6. COUNTRY: USA
 7. ZIP: 02109
 8. COMPUTER READABLE FORM:
 9. MEDIUM TYPE: Diskette
 10. COMPUTER: IBM Compatible
 11. OPERATING SYSTEM: DOS
 12. SOFTWARE: FASTSEQ Version 1.5
 13. CURRENT APPLICATION DATA:
 14. APPLICANT: K. NEMER
 15. FILING DATE: 02/29/1995
 16. CLASSIFICATION: 536
 17. PRIOR APPLICATION NUMBER: 08/254,494
 18. FILING DATE: 06-JUN-1994
 19. PRIOR APPLICATION DATA:
 20. APPLICATION NUMBER: JP-079996-1991
 21. FILING DATE: 12-APR-1991
 22. PRIOR APPLICATION DATA:
 23. APPLICATION NUMBER: JP-085396-1991
 24. FILING DATE: 14-APR-1991
 25. PRIOR APPLICATION DATA:
 26. APPLICATION NUMBER: JP 022321-1992
 27. FILING DATE: 07-FEB-1994
 28. ATTORNEY/AGENT INFORMATION:
 29. NAME: RESNICK, DAVID S.
 30. REGISTRATION NUMBER: 41777
 31. REFERENCE/DOCKET NUMBER: 41777
 32. TELEPHONE: (617) 523-3400
 33. TELEFAX: (617) 523-6440
 34. TELEX: 200291 STEE UR
 35. INFORMATION FOR SEQ ID NO: 3:
 36. LENGTH: 1120 base pairs
 37. TYPE: nucleic acid
 38. STRANDEDNESS: double
 39. TOPOLOGY: linear
 40. MOLECULE TYPE: cDNA to genomic RNA
 41. HYPOTHETICAL: NO
 42. ANTI-SENSE: NO
 43. FRAGMENT TYPE:
 44. ORIGIN: human
 45. ORGANISM: human
 46. CELL TYPE: breast carcinoma
 47. CELL LINE: ZR-75-1
 48. NAME/KEY: Coding Sequence
 49. LOCATION: 112...796
 50. OTHER INFORMATION: E Mat peptide
 51. us-09-727-770-2

Alignment Scores:
 Pred. No. Length: 1120
 Score: 100.50 Matches: 40
 Percent Similarity: 44.68% Conservat: 23
 Best Local Similarity: 27.97% Mismatches: 46
 Query Match: 9.52% Indels: 44
 Gaps: 7

us-09-727-770-2 (1-205) x us-08-408-2226-3 (1-1120)

07 83 AsnAspAspAsuSerLeuTyrSerSerProLeuLeuLeuLeuLeuLeuLeu 99
 24 242 AATAAATTCAGATCTTCAACAGAGAGTCAATTTG-----ATC 304
 07 100 Aa 119
 06 404 GAG 363

07 120 ProAlaLeuPheValGlyPheA-----133
 06 464 TCCAGTGTAAAGTGGGAGTCTCTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCT 423
 07 134 -----LysValLeuLeuLeuSerThrLysGlnProLeu 144
 06 424 GCTGGGGGAACTGGGGGATATTTCACAAAGGATATGATGATTAAAGAAATCCAGAGTTT 483
 07 145 SerLysPheThrSerHisAAlaGlyGlyLysAspHisProSerPheSerPhe 164
 06 484 TACAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 543
 07 164 sGLPFGAspThrHisValLeuGlySerTyrProSerValVa 178
 06 544 G---CAATGACTATGGTTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 505
 07 178 TAspLeuLeuSerValGlyGlySerPheAValHisAlaLeuLeuAAlaAPr 198
 06 590 ----TAAATCAGACATCTGAGGCAAGA-----AGAGAGATCTGAGAAAGG 645
 07 198 uproPro 200
 06 646 AAGTCT 652

RESULT 7

us-08-254-493-3

1. Sequence 3, Application 05/08/254493
 2. Patent No. 5439886
 3. GENERAL INFORMATION:
 4. APPLICANT: KIKUYAMA, SHUICHI
 5. APPLICANT: KUYAMA, MASARU
 6. APPLICANT: KUYAMA, MASARU
 7. APPLICANT: SENOO, MASAHARU
 8. TITLE OF INVENTION: PERFECTURAL ANTIBODY, POLYPEPTIDES AND
 9. NUMBER OF SEQUENCES: 7
 10. CORRESPONDENCE ADDRESS:
 11. ADDRESSEE: DAVID S. RESNICK, FIRM, RESEKIN, EXPERTS & CONSULTANTS
 12. STREET: 140 WATER STREET
 13. CITY: BOSTON
 14. STATE: MASSACHUSETTS
 15. COUNTRY: US
 16. ZIP: 02109
 17. COMPUTER READABLE FORM:
 18. MEDIUM TYPE: Floppy disk
 19. COMPUTER: IBM PC Compatible
 20. OPERATING SYSTEM: PC DOS/MS DOS
 21. SOFTWARE: Patent In Release #1.0, Version #1.25
 22. CURRENT APPLICATION DATA:
 23. APPLICATION NUMBER: 08/38/254,493
 24. FILING DATE:
 25. CLASSIFICATION: 435
 26. PRIOR APPLICATION DATA:
 27. APPLICATION NUMBER: 08/07/965552
 28. FILING DATE: 09-APR-1992
 29. PRIOR APPLICATION DATA:
 30. APPLICATION NUMBER: JP 079996-1991
 31. FILING DATE: 12-APR-1991
 32. PRIOR APPLICATION DATA:
 33. APPLICATION NUMBER: JP 085396-1991
 34. FILING DATE: 17-APR-1991
 35. PRIOR APPLICATION DATA:
 36. APPLICATION NUMBER: JP 022321-1992
 37. FILING DATE: 07-FEB-1992
 38. ATTORNEY/AGENT INFORMATION:
 39. NAME: RESNICK, DAVID S.
 40. REGISTRATION NUMBER: 41777
 41. REFERENCE/DOCKET NUMBER: 41777
 42. TELEPHONE: (617) 523-3400
 43. TELEFAX: (617) 523-6440
 44. TELEX: 200291 STEE UR
 45. INFORMATION FOR SEQ ID NO: 3:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; HYPOTHEICAL: NO
; ANTI SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HUMAN
; CELL TYPE: BREAST CARCINOMA
; CELL LINE: ZR-75-1
; FEATURE:
; NAME/KEY: ccs
; LOCATION: 112...795
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 115...795
; US-08-254-493-3

```

Alignment Scores:

```

Pred. No.: 0.045 Length: 1120
Score: 96.50 Matches: 45
Percent Similarity: 43.36% Conservat: 22
Best Local Similarity: 27.97% Mismatches: 47
Query Match: 9.14% Indels: 34
DB: 1 Gaps: 7

```

US-09-727-770-2 (1-205) x US-08-254-493-3 (1-1120)

```

QY 80 AsuAspAsnSerLeuIyTSerSerPheLeuGlnGlyAlaGlyLeuSerGlyLeu 99
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 262 AUAATAATATCCAGCTCTACACAGGAGTCTATATCTG-----ATC 303
QY 100 AlaAlaGlyPheAlaIleValIleValGlyAspThrGlyGlyGlyGlyPheAlaGlnGln 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 304 GAGCGCGCGCGCCCTCAAGATGCTGGTGGGCTTCCTCGGAGTCTATATAGGAGAG 363
QY 120 ProArgLeuPheValGlyMet-----IleLeuIleLeuIlePheAla----- 133
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 364 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
QY 134 -----LysValIleLeuLeuSerThrLysGlnProLeu 144
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 424 GTCGCGCGCATCTGGGATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
QY 145 SerLysProPheSerHisAlaGlyAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 161
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 484 TACAAGACACACCTACACACACACACACACACACACACACACACACACACACACACAC 543
QY 164 SGIProAspThrHis ValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 179
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 544 G--CGATCGACTATGCTTCAACTGCTGTGGTTTGGCTGGGCGGTGGAACAGT----- 595
QY 178 LaspLeuLeuSerValGlyGlySerAlaValHisAspGlnGlnAlaArgPheGlyProLe 198
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 596 -----TATCTCAGACATCTGCCCAAGA-----AGAGATATCTGAACCTTCACGCTG 645
QY 198 uProPro 200
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 646 AAGTCCT 652

```

RESULT 6

US-08-816-755-1

Sequence 1, Application US/08816755

Patent No. 5847508

GENERAL INFORMATION:

APPLICANT: Arteth, Anthony

APPLICANT: Shabon, Osman

TITLE OF INVENTION: No. 5847508el Membrane Type Matrix

NUMBER OF SEQUENCES: 6

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,755
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 4.45
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 44,444
; PREFERENCE RIGHT NUMBER: CH50007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-816-755-1

```

```

Alignment Scores:
Pred. No.: 2.01 Length: 4807
Score: 89.50 Matches: 54
Percent Similarity: 39.51% Conservat: 27
Best Local Similarity: 26.44% Mismatches: 56
Query Match: 8.48% Indels: 68
DB: 2 Gaps: 12

```

US-09-727-770-2 (1-205) x US-08-816-755-1 (1-4807)

```

QY 20 ThrMetValSerSerGlyLeuGlyAlaAlaGlyGlyMetAlaLysAsnGlyThr----- 47
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2905 ACCCTGCGTCTCTGCGATTCGGGTGCGCTTACGCTGAGACACAGACACCTCTGCT 2964
QY 49 -----GlyIleMetAlaMetSerValMetThrProGlnIleuIleHisMetLysSer 54
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2465 GCGCTGGGAGCTGCTTAAGCAAAATCTCTGTTCCTACAGATGGTCACTGGTCTCC 604
QY 55 IletProValValMetAlaGlyIleIleIleIleIleGlyLeuValAlaValPro 74
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4025 GTTCTT----- 4036
QY 75 ProAlaSerLeuAsnAspAspAsnSerLeuGlySerSerPheLeuGlnIleuGlyAla 94
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3037 TGTTCATCATCTCTG----- 3072
QY 95 GlyLeuSerGlyLeuAlaAlaGlyPheAlaIleValIleValGlyAspThrGlyLysGly 114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3073 GGGAGG----- 3099
QY 115 GlyThrAlaGlnIleProArgLeu-----PheValIleMetIleLeuIleLeuIle 131
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 3097 -----GAACAATTCGGAGATATTAGTGAATCAAGTTCTGACAGCTTTTAA 147
QY 142 PheAlaLysValLeuIle-----LeuSerThrLysGlnProGlnSerLysProThr 148
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4148 TTGCAAAAGGACCTTTATTAGCTACACCTGTCCTACATCAAGAACTGCTTAAAGGAACT 4207

```



```

: ATTORNEY/AGENT INFORMATION:
: NAME: Brocke, A. Andre
: REGISTRATION NUMBER: 46,474
: REFERENCE/DOCKET NUMBER: 1640P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309,8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 36:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 21706 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-961-527-36
:
: Alignment Scores:
: Pred. No.: 132 Length: 21706
: Score: 84.00 Matches: 48
: Percent Similarity: 46.59% Conservations: 25
: Best Local Similarity: 49.57% Mismatches: 44
: Query Match: 7.95% Indels: 40
: Gaps: 10
:
: US-09-727-770-2 (1 205) x US-08-961-527-36 (1-21706)
:
: QY 23 MetValSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValValM 60
: DB 20711 GATATGCGGGGATGATGACCGGACGCGGAGATGATTCATATTCACGCGGG 20656
:
: QY 40 MetValSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValValM 60
: DB 20655 -----ATGTCATCATGAGAGATGATTCAGAT-----GGTGTATACAGATGTAG 20607
:
: QY 60 MetValSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValValM 71
: DB 20606 GATATGCGGGGATGATGACCGGACGCGGAGATGATTCATATTCACGCGGG 20547
:
: QY 71 MetValSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValValM 85
: DB 20546 GTGGCATCATGATTCATGATCAAGAGAGATGATTCATCAAGAGATGATGATG 20487
:
: QY 85 -----TyrSerSerPheLeuGluGlyAlaGlyProSerGluValAlaGly 102
: DB 20486 GATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20427
:
: QY 103 -----PheAlaIleValIleValGlyAspThrGlyLysCysGlyThrAlaGln 119
: DB 20426 CAATGACGCTTTCATATGCTTT-----GTAATATCTGCGCGTCAAG 20379
:
: QY 119 LeuProAlaLeuProValGlyMet IleLeuIleLeuIle----- 131
: DB 20478 ATACGATATGTTTATTACTATGATGATGATGATGATGATGATGATGATGAT 20419
:
: QY 132 --PheAlaValValLeuLeuSerThrLysGlnProValSerLys 146
: DB 20318 TCTTCAAGCAACTCTTGATTAACCTTTCTTGATGAACTCTTGCTTAA 20272
:
: RESULT 14
: US-08-801-344-7
: Sequence 7, Application US/0880144
: Patent No. 6,957,149
: GENERAL INFORMATION:
: APPLICANT: Cameron, Douglas C.
: APPLICANT: Shaw, Anita J.
: APPLICANT: Attaras, Nedim B.
: TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPYDIOL
: NUMBER OF INVENTION: FROM SUGAR
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lemitt Ross & Stevens, S.C.
: STREET: 8000 Excelsior Drive, Suite 401
: CITY: Madison
: STATE: WI

```

```

: COUNTRY: U.S.A.
: ZIP: 53717 1914
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC DOS/MS DOS
: SOFTWARE: Patent In Release #1.0, Version #1.40
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 05/98/851,444
: FILING DATE:
: CLASSIFICATION: 445
: ATTORNEY/AGENT INFORMATION:
: NAME: Sara, Charles S.
: REFERENCE/DOCKET NUMBER: 09823,017
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 608 841 2100
: TELEFAX: 608-821-2106
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3524 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: NUCLEIC ACID TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ART: GENE, NO
: ORIGINAL SOURCE:
: ORGANISM: Pyridine nucleotide Transhydrogenase gene
: US-08-801-344-7
:
: Alignment Scores:
: Pred. No.: 16,1 Length: 3524
: Score: 81.50 Matches: 44
: Percent Similarity: 43.14% Conservations: 24
: Best Local Similarity: 23.94% Mismatches: 47
: Query Match: 7.72% Indels: 48
: Gaps: 5
:
: US-09-727-770-2 (1 205) x US-08-801-344-7 (1-3524)
:
: QY 18 MetAlaThrMetValSerSerGlyLeuGlyAlaAlaGlySerGlyMetAlaLysAspGlyThr 37
: DB 1627 CTGGGCGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1686
:
: QY 38 Gly-----Ile-MetAla 41
: DB 1687 SCAGGTGAAAGTGAAGGAAAAATGATGATGATGATGATGATGATGATGATGATGAT 1746
:
: QY 41 MetSerValMet-----TrpProGluLeuIleHisMetLysSerIleIleProVal 58
: DB 1747 CTGGCAATCATCTTTTTCGCTGG-----ATGCAAGAGATGATGATGATGATG 1791
:
: QY 58 ValMetAlaGlyIleIleThrIleThrGlyLeuValAlaAlaValProProAlaAsnSer 78
: DB 1792 AGAATGCTTGGCACTTCACCTTTTCGATGATGATGATGATGATGATGATGATGAT 1891
:
: QY 78 LeuAlaAlaSerPheSerLeuGluGluPheProGluLeuLeuValAlaLeuSerIle 98
: DB 1852 GTGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1911
:
: QY 98 ValAlaAlaGlyIlePheAlaIleValIleValIleValIleValIleValIleVal 118
: DB 1912 G-----ATATCTGCTGCGACATGCTTTCAGATGATGATGATGATGATGAT 1991
:
: QY 118 GlnThrGluAlaGluPheValGlyPheIleGluGluIleLeuIlePheAlaValValLeuIleGlu 138
: DB 1952 -----GGGTGGGTTAGTCTTCTTATGTTTATGATGATGATGATGATGAT 1992
:
: QY 138 Ser 149
: DB 1993 CAGC 1996
:
: RESULT 15

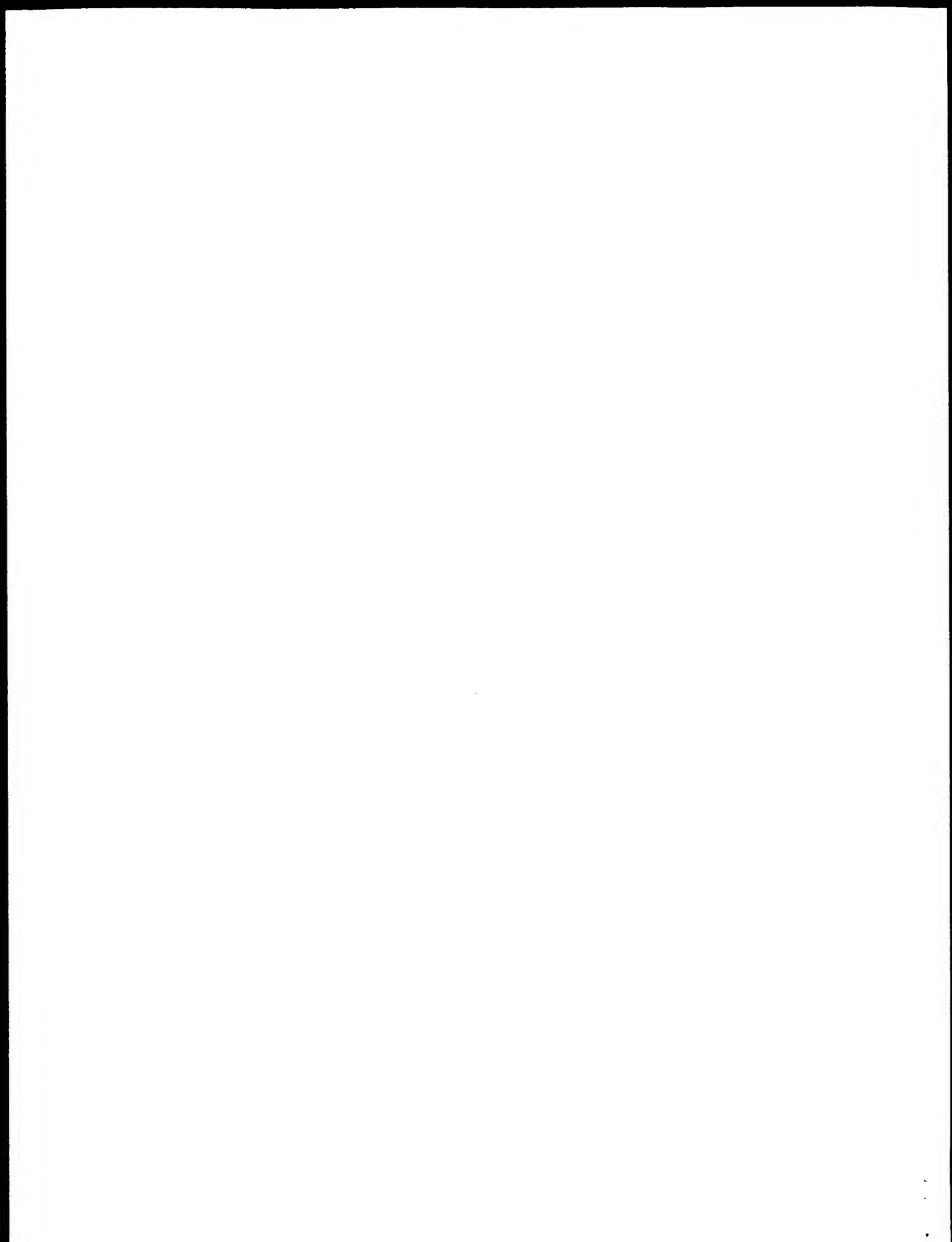
```

US-09-498-599-7
 : Sequence 7, Application US/09498599
 : Patent No. 498452
 : GENERAL INFORMATION:
 : APPLICANT: Cameron, Douglas C.
 : APPLICANT: Shaw, Anita J.
 : APPLICANT: Altaras, Nadia E.
 : TITLE OF INVENTION: MICROBIAL PRODUCTION OF
 : NUMBER OF SEQUENCES: 11
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Dewitt Ross & Stevens S.C.
 : STREET: 8000 Excelsior Drive, Suite 401
 : CITY: Madison
 : STATE: WI
 : COUNTRY: U.S.A.
 : ZIP: 53717-1914
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM pc compatible
 : OPERATING SYSTEM: PC DOS/MS-DOS
 : SOFTWARE: Patent in Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/498,599
 : FILING DATE:
 : CLASSIFICATION:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Sara, Charles S.
 : REFERENCE/DOCKET NUMBER: 09820.037
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 508-831-2100
 : TELEFAX: 508-831-2106
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 3524 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : Molecule TYPE: DNA (genomic)
 : HYDROLYTICAL: NO
 : ANTI-SENSE: NO
 : ORIGINAL SOURCE:
 : ORGANISM: Pyridine nucleotide transhydrogenase gene
 : US-09-498-599-7

Search completed: December 28, 2002, 12:21:11
 Job time : 87 secs

Alignment Scores:
 Pred. No.: 16,1 Genoth: 3524
 Score: 81.50 Matches: 34
 Percent Similarity: 40.14% Conservative: 24
 Best Local Similarity: 23.94% Mismatches: 47
 Query Match: 7.72% Indels: 38
 Gaps: 5

US-09-727-770-2 (1-205) x US-09-498-599-7 (1-3524)
 QY 18 MetAlaThrMetValSerGorGlyLeuAlaAlaTysGlyMetAlaLysAsnGlyThr 37
 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572



GenCore version 5.1.1.4
Copyright (c) 1993 - 2002 Computer Ltd.

em nucleic - nucleic search, using sw model

Run on: December 28, 2002, 09:42:06 : Search time 2579 Seconds
(without alignments)
7143.108 Million cell updates/sec
Title: US-09-727-770-1
Perfect score: 633
Sequence: 1 teacacacatgataa.....ctatggaat-ttactaac 633

Scoring table: IDENTITY_NUC
Gapop 10.0 : Gapext 1.0

Sequences: 2054440 seqs 1455402478 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum hit seq length: 0
Maximum hit seq length: 20000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Genbank:*

- 1: db_ba:*
- 2: db_btq:*
- 3: db_in:*
- 4: db_m:*
- 5: db_ow:*
- 6: db_pat:*
- 7: db_ph:*
- 8: db_pl:*
- 9: db_pr:*
- 10: db_ro:*
- 11: db_sts:*
- 12: db_sy:*
- 13: db_un:*
- 14: db_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hab:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_un:*
- 28: em_vl:*
- 29: em_btq_ham:*
- 30: em_btq_hay:*
- 31: em_btq_ether:*
- 32: em_btq_mos:*
- 33: em_btq_pln:*
- 34: em_btq_pln:*
- 35: em_btq_rnd:*
- 36: em_btq_mam:*
- 37: em_btq_vrt:*
- 38: em_sy:*
- 39: em_haq_hum:*
- 40: em_haq_mus:*
- 41: em_haq_other:*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	633	100.0	633	6	AX337957	AX337957 Sequence
2	626	98.9	633	6	AX337959	AX337959 Sequence
3	626	98.9	151999	2	AC005973	AC005973 Homo sapi
4	626	98.9	184595	2	AC015913	AC015913 Homo sapi
5	615	97.2	149384	2	AC104333	AC104333 Homo sapi
6	375	59.2	1086	9	BC004547	BC004547 Homo sapi
7	375	59.2	1122	9	BC009290	BC009290 Homo sapi
8	375	59.2	1143	9	BC007489	BC007489 Homo sapi
9	473.4	59.0	1078	9	BC007759	BC007759 Homo sapi
10	361	57.0	1162	9	BUMPCH20CA	M62762 Human vacuo
11	332	52.4	1975	9	BC010137	BC010137 H. sapi
12	317.6	50.2	110900	9	HS1392M17	AL049843 Human DNA
13	304.8	48.2	1129	10	M001342	U33842 Mus musculu
14	304.8	48.2	1153	10	M08WVP	M64298 Mouse vacuo
15	294.4	46.5	1082	10	RATSDHATP	D10874 Rattus norv
16	290.5	45.9	1133	4	B_VAT1952	U04826 Rattus norv
17	289.2	45.7	266574	2	AC087129	AC087129 Mus muscu
18	253.9	40.1	146571	2	AC129276	AC129276 Rattus no
19	245.8	38.8	469	4	AF027795	AF027795 Rattus norv
20	214	33.8	14283	2	AC093525	AC093525 Homo sapi
21	212.6	33.5	210105	2	AC125419	AC125419 Rattus no
22	190.2	30.0	902	5	AY099523	AY099523 Rattus norv
23	192.8	28.9	131785	2	AC099795	AC099795 Homo sapi
24	192.8	28.9	192180	2	AC023331	AC023331 Homo sapi
25	175.2	27.7	616	5	U01520AM	X52002 T. maritima
26	173.2	27.4	186236	2	AC116716	AC116716 Mus muscu
27	171.6	27.1	13813	10	AR059662	AC059662 Mus muscu
28	164.8	26.0	111539	2	AC098526	AC098526 Rattus no
29	157.2	24.9	1139	3	GHVHATP	X55979 Prosophila
30	143.2	22.6	1429	3	KS5ATP153	X55651 M. scab. mRN
31	139.4	22.0	518	3	AF277150	AF277150 Selenopel
32	137	21.6	854	3	AF008924	AF008924 Aedes aeg
33	136.2	21.5	918	3	AF008924	AF008924 Aedes aeg
34	135.2	21.4	22979	2	AC012221	AC012221 Prosophil
35	135.2	21.4	83340	3	AC004438	AC004438 Prosophil
36	135.2	21.4	159928	3	AC010565	AC010565 Prosophil
37	135.2	21.4	186653	3	AC010043	AC010043 Prosophil
38	125.2	21.4	281869	3	AC003544	AC003544 Prosophil
39	133.6	21.1	132182	9	AC092916	AC092916 Homo sapi
40	123.6	21.1	196640	2	AC026099	AC026099 Homo sapi
41	132	20.9	1015	3	H01ATPASEV	L11991 Helicobact
42	122.4	19.3	6186	10	AF356008	AF356008 Mus muscu
43	114	18.0	102413	2	AC020924	AC020924 Prosophil
44	114	18.0	154985	3	AC004641	AC004641 Prosophil
45	114	18.0	179283	3	AC099360	AC099360 Prosophil

ALIGNMENTS

RESULT 1	AX337957	Sequence 1 from Patent	633 bp	DNA	Linear	PAI 09 JAN 2002
LOCUS	AX337957	Sequence 1 from Patent	633 bp	DNA	Linear	PAI 09 JAN 2002
DEFINITION	AX337957	Sequence 1 from Patent	633 bp	DNA	Linear	PAI 09 JAN 2002
ACCESSION	AX337957	Sequence 1 from Patent	633 bp	DNA	Linear	PAI 09 JAN 2002
VERSION	AX337957.1	GI:18128674				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE	Li, Z., Wei, M.H., Ketchum, K. and Beasley, E.M.					
	Isolated human transporter proteins, nucleic acid molecules					
	encoding human transporter proteins, and uses thereof					

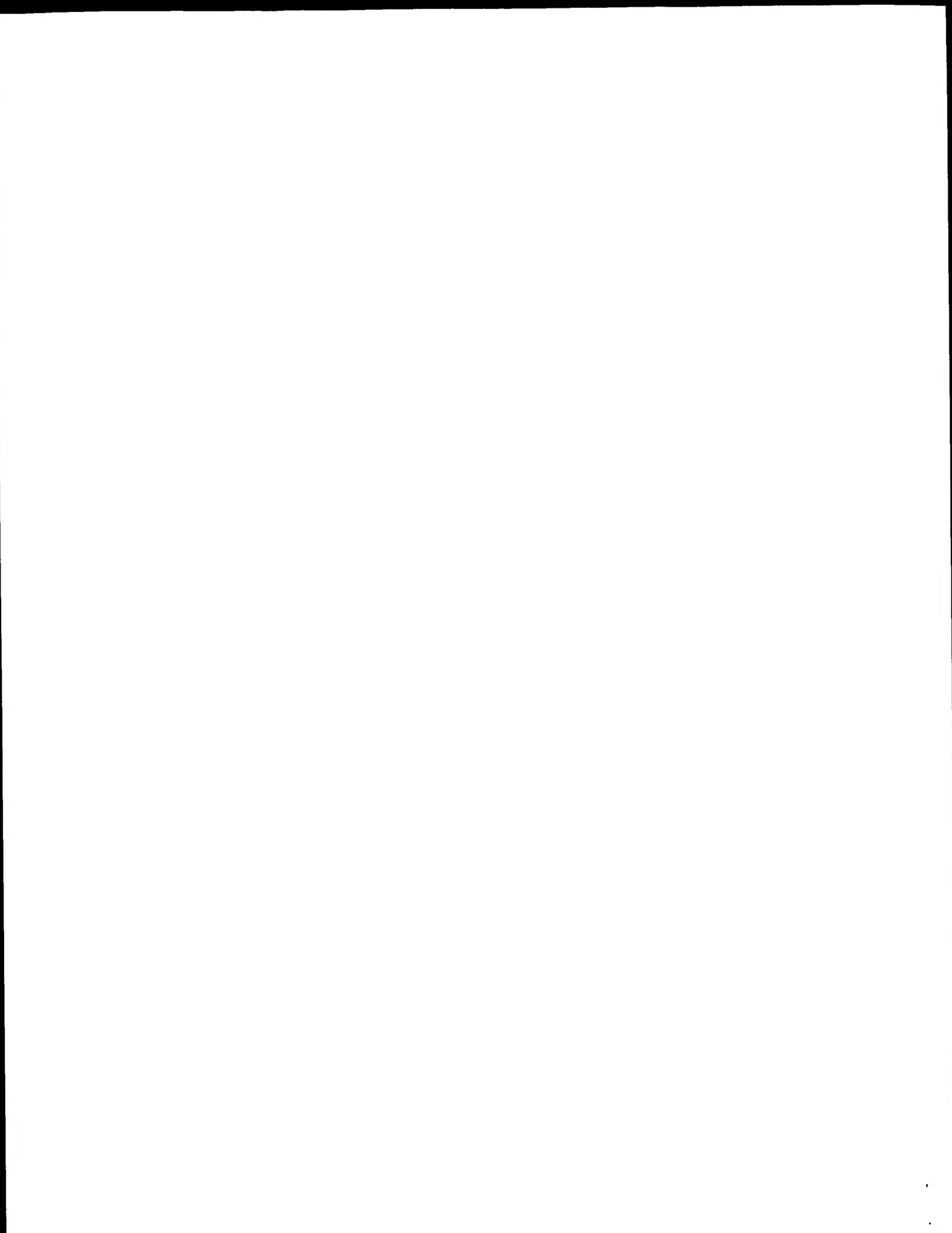
[illegible]

HMMPCHSU0A		HMMPCHSU0A		1162 bp	mRNA	F13-ATP17-JAR	99%
Accession:		Human vacuolar H(+)-ATPase proton channel subunit mRNA, complete cds.					
M62762							
M62762.1	G1:189675						
vacuolar H(+)-ATPase,							
Homo sapiens (tissue library: cDNA) cDNA to mRNA.							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;							
1 (bases 1 to 1162)							
Gillespie,G.A., Somlo,S., Germino,G.G., Weinstein,Saslow,D., and							
Reeders,S.P.							
Cpg island in the region of an autosomal dominant polycystic kidney							
disease locus defines the 5' end of a gene encoding a putative							
proton channel							
Proc. Natl. Acad. Sci. U.S.A. 88 (18), 4289-4293 (1991)							
91249553							
1709739							
Location/Qualifiers							
1..1162							
/organism="Homo sapiens"							
/db_xref="taxon:9606"							
/map="16p13.4"							
/cell_type="Bcell"							
/tissue="Lib_mRNA"							
1..1162							
/gene="ATP11"							
231..698							
/gene="ATP11"							
/cdon_start=1							
/product="vacuolar H+ ATPase proton channel subunit"							
/protein_id="AA60089.1"							
/db_xref="GI:189676"							
/db_xref="GDB:660-128-141"							
7133513:138 "MSPSPSGYASPFAYMGSAAMVFSAAGAAATKSGTGLAAM							
SWMKPEQPKSLIPVVMASITAIYLIVAVLIARIGLSTFELSTALGGATGVSRIS							
GLAAGQILGVLDVAGVGAQQPRLLVGMILLTEAKVLAVGLIIVALIISTR"							
BASE COUNT	176 a	421 c	412 g	253 t			
ORIGIN	chromosome 16, map position p14.4.						
Query Match	97.0%; Score: 661; Dbs 9; Length 1162;						
Best Local Similarity	78.1%; Prod. No.: 8,5e-80;						
Matches	516; conservative	0; Mismatches	100; Indels	4%; Gaps			
QY	12	GTCTCAACAGCGCCGAGTAGTAAGTTTGTTTTGCATTTCACACTTCAGAGGTATGGTATATAT	71				
		TTTTTTT	TT				
DB	239	GTCACAGAGCGGCCGAGTAGTAAGTTTGTTTTTCGCTCTCATAGRGCTCTCTGCTATATAT	298				
QY	72	GGTCTCAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	131				
		TT	TT				
DB	299	GGCTCTCAGCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	358				
QY	132	CATGCTGCTGATATGGCGCAGAGTCTCATCCATCAAGTCTATCATCCAGTGGTATGGT	191				
		TT	TT				
DB	359	CATGCTGCTATGCGCGCGAGTACAT - CATGAAGCTATCATCCAGTGGTATGGT	415				
QY	192	TGGTATCATCACATCTATATGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	251				
		TT	TT				
DB	416	TGGATCATCGGCATCTATGRCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	475				
QY	252	TGCAAAAGTGTCTATATGTAAGTTTCTCTCATCTGGCGT	299				
		TT	TT				
DB	476	CGATCATCGCTCTACCAAGATCTCTCTCATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	535				
QY	400	TGGTCTGGACGCGGCTTTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	459				
		TT	TT				
DB	536	CGGCTGCGACGCGCTTTCATCATGCTATCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGG	595				
QY	360	CGACGAGCGCGGACTATTTCGAGGATATATATATGATCTTCTTTCTTTCTTTCTTTCTTTCT	414				
		TT	TT				
DB	596	CGACGAGCGCGGACTATTTCGAGGATATATATATGATCTTCTTTCTTTCTTTCTTTCTTTCT	655				

[illegible]

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Hominidae: Homo.
 1 (bases 1 to 110900)
 Collier, R.
 Direct Submission
 Submitted (15-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail: r.collier@sanger.ac.uk
 CH10
 On Sep 6, 1999 this sequence version replaced g1:5738452.
 This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 EMBL: SW; SWISSPROT: Tr; TrEMBL: Wp; WORMPEP: Information
 on the WORMPEP database can be found at
<http://www.sanger.ac.uk/ftp/seqs/1.0-rnaseq/wormpep> This sequence
 was generated from part of bacterial clone contains of human
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/ftp/seqs/1.0-rnaseq/>
 RP3-392M17 is from the library p991-3 constructed at the Roswell
 Park Cancer Institute by the group of Pieter de Jong. For further
 details see <http://zookey.med.toronto.edu/>
 VECTOR: pCYPAC2
 This sequence is the entire insert of clone RP3-392M17.
 Location/Qualifiers
 1..110900
 /organism="Homo sapiens"
 /db_xref="taxon:6066"
 /chromosome="6"
 /map="p12-4-21.2"
 /clone="RP3-392M17"
 /clone_1fb="RP3-392M17-3"
 repeat_region
 1..208
 /note="AluX repeat: matches 1..197 of consensus"
 repeat_region
 403..709
 /note="AluX repeat: matches 1..310 of consensus"
 repeat_region
 751..939
 /note="AluX repeat: matches 6..196 of consensus"
 repeat_region
 956..1133
 /note="L1MA2 repeat: matches 6058..6068 of consensus"
 repeat_region
 1194..1495
 /note="AluX repeat: matches 1..299 of consensus"
 repeat_region
 1496..1520
 /note="L1MA2 repeat: matches 6044..6068 of consensus"
 repeat_region
 1521..1812
 /note="AluX repeat: matches 1..293 of consensus"
 repeat_region
 1813..2737
 /note="L1MA2 repeat: matches 514..6044 of consensus"
 repeat_region
 2739..2917
 /note="AluX repeat: matches 112..308 of consensus"
 repeat_region
 2918..3211
 /note="AluX repeat: matches 1..292 of consensus"
 repeat_region
 3213..3272
 /note="L1MA2 repeat: matches 5086..5142 of consensus"
 repeat_region
 3273..3405
 /note="FRAM repeat: matches 91..40 of consensus"
 repeat_region
 3465..3762
 /note="AluX repeat: matches 3..299 of consensus"
 repeat_region
 3816..3864
 /note="L1P3 repeat: matches 360..404 of consensus"
 repeat_region
 3898..4024
 /note="L1P4 repeat: matches 48..201 of consensus"
 repeat_region
 4223..4532

/note "AluX repeat: matches 1..307 of consensus"
 4665..4967
 /note "AluX repeat: matches 1..302 of consensus"
 4977..5278
 /note "AluX repeat: matches 1..294 of consensus"
 5414..6348
 /note "L1P3 repeat: matches 1..595 of consensus"
 6372..6661
 /note "AluX repeat: matches 3..293 of consensus"
 6696..7000
 /note "AluX repeat: matches 1..306 of consensus"
 7230..7435
 /note "L1MA2 repeat: matches 684..902 of consensus"
 7686..7906
 /note "AluX repeat: matches 9..240 of consensus"
 7967..8100
 /note "AluX repeat: matches 1..134 of consensus"
 8101..8401
 /note "AluX repeat: matches 1..301 of consensus"
 8402..8577
 /note "AluX repeat: matches 134..309 of consensus"
 8605
 8617..8803
 /note "AluX repeat: matches 1..303 of consensus"
 8904..9072
 /note "L1MA2 repeat: matches 2652..2808 of consensus"
 9101..9164
 /note "AluX repeat: matches 75..134 of consensus"
 9218..9514
 /note "AluX repeat: matches 1..302 of consensus"
 9789..10092
 /note "AluX repeat: matches 1..311 of consensus"
 10159..10338
 /note "L1MA2 repeat: matches 189..370 of consensus"
 10380..10687
 /note "AluX repeat: matches 1..303 of consensus"
 10690..10841
 /note "76 copies 2 mer 11-69 conserved"
 10845..11158
 /note "AluX repeat: matches 1..211 of consensus"
 11160..11263
 /note "72 copies 2 mer 10-69 conserved"
 11267..11576
 /note "AluX repeat: matches 1..302 of consensus"
 11577..11612
 /note "18 copies 2 mer 10-68 conserved"
 11645..11902
 /note "AluX repeat: matches 1..259 of consensus"
 12362..13027
 /note "CpG island"
 /note "Evidence not experimental"
 complement (12455..12921)
 /note "33392H7.1"
 complement (12455..12921)
 /note "d0492M17.1"
 /note "d0492M17.1 (similar to AluX (dAluX), H-
 transporting lysosomal (vacuolar proton pump) 16kDa);
 similar to AluX (dAluX), H-transporting lysosomal
 (vacuolar proton pump) 16kDa; match: ERAS = E0AF92709;
 match: proteins: SW:018882 SW:024956 SW:024449"
 /pseudo
 /codon_start 1
 /evidence not experimental
 14274..14592
 /note "AluX repeat: matches 1..303 of consensus"
 14689..14810
 /note "FRAM repeat: matches 1..121 of consensus"
 14915..15043
 /note "L1MA2 repeat: matches 1..129 of consensus"
 15057..15615



XX Li Z, Wei M, Ketchum K, Beasley EM;
XX WPI: 2002-566438/60.
XX p-PSDB: AA015455.
XX New human transporter proteins related to the proton ATPase transporter
XX subfamily for diagnosing and treating diseases mediated by the
XX transporter protein and for identifying modulators
XX Claim 23; Fig 1; 70pp; English.
XX The invention comprises the amino acid and coding sequence of a human
XX transporter protein. Transporter proteins regulate many different
XX functions of a cell, including cell proliferation, differentiation and
XX signalling processes. Transporter proteins function by regulating the
XX flow of molecules such as ions and macromolecules into and out of cells.
XX The gene for the human transporter protein is located on chromosome 17.
XX The DNA and protein sequences of the invention are useful for identifying
XX agents that modulate or bind to the human transporter protein. The DNA
XX and protein sequences are useful for treating a disease or condition that
XX is mediated by a human transporter protein. The DNA and protein sequences
XX of the invention may also be used in the construction of transgenic
XX animals, pharmacogenomic analysis and for tissue typing. The present DNA
XX sequence encodes the human transporter protein of the invention.
XX Sequence: 633 bp; 128 A; 202 C; 153 G; 150 T; 0 other;
SQ
Query Match 100.0%; Score 633; DB 24; Length 633;
Best Local Similarity 100.0%; Pred. No. 6, 8e 176;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAGAACATGCGCAACACAGACCGGACATGCTTTCATCAGATCTCGGCTCT 60
DB 1 TCACAGAACATGCGCAACACAGACCGGACATGCTTTCATCAGATCTCGGCTCT 60
QY 61 ATGCGACATGCTTTCATCAGATCTGCTTTCATGCGGACATGCGGACATGCGGAC 120
DB 61 ATGCGACATGCTTTCATCAGATCTGCTTTCATGCGGACATGCGGACATGCGGAC 120
QY 121 GCGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 180
DB 121 GCGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 180
QY 181 GCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 240
DB 181 GCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 240
QY 241 TCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 300
DB 241 TCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 300
QY 301 GCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 360
DB 301 GCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 360
QY 361 GCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 420
DB 361 GCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 420
QY 421 CTGCGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 480
DB 421 CTGCGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 480
QY 481 GCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 540
DB 481 GCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 540
QY 541 GTTGAATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 600
DB 541 GTTGAATGCGGACATGCGGACATGCGGACATGCGGACATGCGGACATGCGGAC 600
QY 601 CTCGCGCGCGCATGCTGCTGCGACATGCTGCAATGCAATGCAATGCAATGCAAT 633

DB 601 CTCGCGCGCGCATGCTGCTGCGACATGCTGCAATGCAATGCAATGCAATGCAAT 633
RESULT 2
AAL44086
ID AAL44086 standard: DNA: 6339 bp.
XX AAL44086;
XX 03-OCT-2002 (first entry)
XX Human transporter protein gene sequence.
XX Human: gene, ds; gene therapy; transporter protein; cell proliferation;
XX cell differentiation; cell signalling; chromosome 17; transgenic animal;
XX transporter protein-mediated disease; pharmacogenomic analysis;
XX tissue typing.
XX Homo sapiens.
XX key location/qualifiers
XX variation /tag a replace (559, A)
XX FT /standard_name "Single nucleotide polymorphism"
XX FT /note "The present base is optionally absent"
XX FT 3000..3617
XX FT /tag b /product "Human transporter protein"
XX FT replace (4648, C)
XX FT /standard_name "Single nucleotide polymorphism"
XX FT replace (5446, T/C)
XX FT /tag d /standard_name "Single nucleotide polymorphism"
XX FT replace (5808, G)
XX FT /tag e /standard_name "Single nucleotide polymorphism"
XX FT replace (5892, C)
XX FT /tag f /standard_name "Single nucleotide polymorphism"
XX FT replace (6071, G)
XX FT /tag g /standard_name "Single nucleotide polymorphism"
XX W0200194380-A2.
XX PN
XX 13-DEC-2001.
XX 31-MAY-2001; 2001WO-0517511.
XX 02 JUN 2000; 2000US-2688560.
XX 04 DEC 2000; 2000US-0727770.
XX (APPL-) APPLERA CORP.
XX Li Z, Wei M, Ketchum K, Beasley EM;
XX WPI: 2002-566438/60.
XX p-PSDB: AA015455.
XX New human transporter proteins related to the proton ATPase transporter
XX subfamily for diagnosing and treating diseases mediated by the
XX transporter protein and for identifying modulators
XX Claim 23; Fig 1; 70pp; English.
XX The invention comprises the amino acid and coding sequence of a human
XX transporter protein. Transporter proteins regulate many different
XX functions of a cell, including cell proliferation, differentiation and
XX signalling processes. Transporter proteins function by regulating the
XX flow of molecules such as ions and macromolecules into and out of cells.
XX The gene for the human transporter protein is located on chromosome 17.
XX The gene for the human transporter protein is located on chromosome 17.

Tue Dec 31 14:03:11 2002

analyzing gene expression in human bone marrow
 Example 4. SEQ ID NO: 14176; 656bp. Sequence Listing: English.
 The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukemia and myeloma. The present sequence is one of the probes of the invention.

Sequence 612 BP; 120 A; 195 C; 151 G; 146 T; 0 other.

Query Match: 96.7%; Score 612; DB 22; Length 612;
 Best Local Similarity 100.0%; Prod. No. 9, 96, 170;
 Matches 612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

8 ATATGTCGACAAACAGACGCGGAGTATGCTTTGGTTTTCATCCTGGGTTGATGCTGCA 67
 1 ACATGTCGACAAACAGACGCGGAGTATGCTTTGGTTTTCATCCTGGGTTGATGCTGCA 60
 68 CCATGTCGTCGACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127
 61 CCATGTCGTCGACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 128 TGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 187
 121 TGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 180
 188 TGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 247
 181 TGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 240
 248 ATGATGACAAACAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 307
 241 ATGATGACAAACAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 300
 308 CAGCGGCTTGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 367
 301 CAGCGGCTTGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 360
 368 CCGGACATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 427
 361 CCGGACATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 420
 428 CAAGGACGCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 487
 421 CAAGGACGCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 480
 488 CATTCCGGAACAAACAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 547
 481 CATTCCGGAACAAACAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 540
 548 TTTCAAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 607
 541 TTTCAAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 600
 608 CCGATGCTGTCG 619
 601 CCGATGCTGTCG 612

RESULT 8
 AAT19420
 ID AAT19420 standard; DNA; 612 BP.
 XX
 AC AAT19420;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Probe #9453 for gene expression analysis in human cervical cell samples.
 XX
 KW Probe: human; microarray; gene expression; cervical epithelial cell;

128 TGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 187
 121 TGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 180
 188 TGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 247
 181 TGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 240
 248 ATGATGACAAACAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 307
 241 ATGATGACAAACAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 300
 308 CAGCGGCTTGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 367
 301 CAGCGGCTTGGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 360
 368 CCGGACATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 427
 361 CCGGACATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 420
 428 CAAGGACGCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 487
 421 CAAGGACGCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 480
 488 CATTCCGGAACAAACAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 547
 481 CATTCCGGAACAAACAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 540
 548 TTTCAAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 607
 541 TTTCAAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 600
 608 CCGATGCTGTCG 619
 601 CCGATGCTGTCG 612

RESULT 7
 AAK38619
 ID AAK38619 standard; DNA; 612 BP.
 XX
 AC AAK38619;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed single exon probe SEQ ID NO: 14176.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 XX
 OS microarray; cancer; leukemia; lymphoma; myeloma; ss.
 XX
 PA Homo sapiens.
 XX
 FN W0200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-0500668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX
 PR 03-AUG-2000; 2000US-0632366.
 XX
 PR 21-SEP-2000; 2000US-024697.
 XX
 PR 27-SEP-2000; 2000US-024697.
 XX
 PR 04-OCT-2000; 2000US-0024263.
 XX
 PA (MOLFE) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DK;
 XX
 DE WPI; 2001-488900/53.
 XX
 KW Human genome-derived single exon nucleic acid probes useful for

P1 Penn SG, Hanzel BK, Chen W, Rank DB;
 XX WPI: 2991-476286/51.
 DB Novel single exon nucleic acid probe used to measure gene expression
 XX in a human breast.
 PT Claim 25; SEQ ID No 5136; 422bp; English.
 PS
 XX The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridises at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, staging,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and
 CC non-carcinoma tumours.
 CC Note: the sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WPI
 CC at http://wipo.int/pub/published/pct_sequences.
 XX
 SS Sequence 612 BP; 120 A; 195 C; 151 G; 146 T; 0 other.
 Query Match 96.7%; Score 612; DB 22; Length 612;
 Best Local Similarity 100.0%; Prod. No. 9.9e+170;
 Matches 612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 ACATGTCCAAACACACAGCCGCGAGTATGCTTTGGTTTTCATCATCTCGATGCTATGCTCA 67
 DB 1 ACATGTCCAAACACAGCCGCGAGTATGCTTTGGTTTTCATCATCTCGATGCTATGCTCA 60
 QY 68 CCATGTGCTCCAGTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 127
 DB 61 CCATGTGCTCCAGTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 QY 128 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187
 DB 121 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 188 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 247
 DB 181 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 QY 248 ATGATGACACACAGCTGCTATAGCAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
 DB 241 ATGATGACACACAGCTGCTATAGCAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
 QY 308 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
 DB 301 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 460
 QY 428 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 487
 DB 421 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 488 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547
 DB 481 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 548 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
 DB 541 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 608 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
 DB 601 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612

QY 68 CACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 127
 DB 61 CACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 QY 128 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187
 DB 121 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 188 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 247
 DB 181 TGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 QY 248 ATGATGACACACAGCTGCTATAGCAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
 DB 241 ATGATGACACACAGCTGCTATAGCAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
 QY 308 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
 DB 301 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 460
 QY 428 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 487
 DB 421 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 488 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547
 DB 481 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 548 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 607
 DB 541 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 608 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 619
 DB 601 CAGGCGGCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612

RESULT 10
 AA105145
 ID AA105145 standard; DNA: 612 BP.
 XX
 AC AA105145;
 XX
 DT 09-oct-2001 (first entry)
 XX
 DE Probe #5136 used to measure gene expression in human breast sample.
 XX
 KW probe, human; breast disease; breast cancer; development disorder; ss;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX
 OS Homo sapiens.
 XX
 PR W3200157270-A2.
 XX
 PR 09-aug-2001.
 XX
 PR 29-jan-2001; 2001W0-US00661.
 XX
 PR 04-feb-2000; 2000US-0180312.
 XX
 PR 26-may-2000; 2000US-0207456.
 XX
 PR 30-jun-2000; 2000US-0608408.
 XX
 PR 03-aug-2000; 2000US-0632366.
 XX
 PR 21-sep-2000; 2000US-0244687.
 XX
 PR 27-sep-2000; 2000US-0236359.
 XX
 PR 04-oct-2000; 2000CB-0024263.
 XX
 (MOLECULAR DYNAMICS INC.)

RESULT 11
ABR12666
ID ABR12666 standard; DNA: 612 BP.
XX
AC ABR12666
XX
D1 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe, see human lung cDNA library
XX
XX Human, ds: single exon probe, asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX xeroderma pigmentosum; Gaucher's disease; Niemann-Pick disease;
XX Hirschsprung's disease; carcinoid; carcinoid; pulmonary haemangioma;
XX pulmonary haemangioma; lymphangioleiomyomatosis; Kaposi's sarcoma;
XX primary ciliary dyskinesia; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; primary ciliary dyskinesia; primary ciliary dyskinesia;
XX hyaline membrane disease; open reading frame; ORF.
XX
CS Homo sapiens.
XX
XX W 200146003-A2.
XX
XX 15-NOV-2001.
XX
XX 15-JAN-2001; 2001W-0500665.
XX
XX 04-FEB-2000; 2000US-180121P.
XX 23-MAY-2000; 2000US-207450P.
XX 02-JUN-2000; 2000US-060840P.
XX 01-JUL-2000; 2000US-064236P.
XX 21-SEP-2000; 2000US-244687P.
XX 22-SEP-2000; 2000US-256459P.
XX 04-OCT-2000; 2000CB-0024263.
XX
XX (M-LE-) MOLECULAR DYNAMICS INC.
XX
XX Penn St., Hanzel DK, Chen W, Rank DK;
XX
XX WP1: 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
XX Claim 4: SEQ ID NO.12677 64bp English
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12687 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes, the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung; comparing (a) a subset of the array with
XX a collection of detectable labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the labeled detectable bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon; the probe is included
XX in the above mentioned microarray, assigned exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarray having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 1261 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene

expression analysis, and for identifying exons in a gene, particularly
using human lung derived mRNA and for the study of lung diseases
such as asthma, lung cancer, chronic obstructive pulmonary disease
(COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
Niemann-Pick disease, Hirschsprung's disease, carcinoid, pulmonary
haemangioma, lymphangioleiomyomatosis, lymphangioleiomyomatosis,
pulmonary alveolar proteinosis, Kaposi's sarcoma, fibrocystic
pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
and hyaline membrane disease. The present sequence is a single exon
probe open reading frame of the invention.
Notes: The sequence data for this patent did not form part
of the filed specification. It was obtained in electronic
format directly from WIPRO at
114 W. 4th St., 11th Fl., published patent sequences.
XX
XX Sequence 612 BP, 126 A, 195 G, 161 G, 146 T, 0 other;
XX
XX Quality Math: 96.7%, Score 612, 39 24, Length 612;
XX Best local Similarity: 100.0%; Prod. No. 9, 90-170;
XX Matches 612, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
XX
XX
XX 8 ACATGTCACACAGACGCGGAGTAAGTATGTTTTCACCAATCGGGAGTATGCGCA 67
XX 1 ACATGTCACACAGACGCGGAGTAAGTATGTTTTCACCAATCGGGAGTATGCGCA 60
XX
XX 68 CAATGTCACACAGACGCGGAGTAAGTATGTTTTCACCAATCGGGAGTATGCGCA 127
XX
XX 61 CAATGTCACACAGACGCGGAGTAAGTATGTTTTCACCAATCGGGAGTATGCGCA 120
XX
XX 128 TGGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 147
XX
XX 121 TGGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 180
XX
XX 188 TGGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 247
XX
XX 191 TGGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 240
XX
XX 248 ATGATGACAAAGTCTCTATGACAGTCTCTATGACAGTCTCTATGACAGTCTCTAT 307
XX
XX 241 ATGATGACAAAGTCTCTATGACAGTCTCTATGACAGTCTCTATGACAGTCTCTAT 300
XX
XX 308 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 367
XX
XX 301 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 360
XX
XX 428 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 427
XX
XX 461 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 420
XX
XX 428 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 487
XX
XX 421 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 480
XX
XX 488 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 547
XX
XX 481 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 540
XX
XX 548 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 607
XX
XX 541 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 600
XX
XX 608 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 619
XX
XX 601 CAGCGCGCTTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 612
XX
XX
XX RESULT 12
XX ABR51753
XX ID ABR51753 standard; DNA: 448 BP.
XX
XX AC ABR51753;
XX
XX

DT 01-FEB-2002 (first entry)
 XX Human foetal liver single exon nucleic acid probe #58.
 DE Human foetal liver, gene expression: single exon nucleic acid probe; ss.
 KW Homo sapiens.
 OS Homo sapiens.
 XX W0200157277-A2.
 DN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-0800669.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0642466.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DE;
 XX WPI: 2001-484447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 XX Claim 1: SEQ ID NO 58; 639pp + sequence listing; English.
 PS The invention relates to a single exon nucleic acid probe for
 XX measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at http://wipo.int/pub/published_pat_sequences.
 XX Sequence 448 bp; 79 A; 147 C; 112 G; 110 T; 0 other;
 SQ
 Query Match 59.2%; Score 475; DP 22; Length 448;
 Best local similarity 100.0%; Prod. No. 40-100;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 ACATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 67
 DB 74 ATATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 133
 QY 68 CCAATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 127
 DB 134 CCAATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 193
 QY 128 TGGCCTATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 187
 DB 194 TGGCCTATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 253
 QY 188 TGGCCTATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 247
 DB 254 TGGCCTATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 313
 QY 248 ATGATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 307
 DB 314 ATGATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 373
 QY 408 CAGCGCGCTTTTGGCCTATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCG 367
 DB 474 CAGCGCGCTTTTGGCCTATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCG 433

QY 368 CCGGACATTTTGTAG 482
 DB 434 CCGGACATTTTGTAG 448
 RESULT 13
 ABA21582
 ID ABA21582 standard; DNA; 448 bp.
 XX ABA21582;
 AC ABA21582;
 DT 23-JAN-2002 (first entry)
 XX Probe #48 for gene expression analysis in human heart cell samples.
 DE Human; gene expression, heart, cardiovascular system; predict
 KW cardiovascular disease; hypertension, cardiac arrhythmia;
 KW congenital heart disease; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 XX W0200157274-A2.
 DN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-080066.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0642466.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DE;
 XX WPI: 2001-488899/53.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX Claim 1: SEQ ID NO 48; 530pp; English.
 PS The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC monitoring and prognosticating diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pat_sequences.
 XX Sequence 448 bp; 79 A; 147 C; 112 G; 110 T; 0 other;
 SQ
 Query Match 59.2%; Score 475; DP 22; Length 448;
 Best local similarity 100.0%; Prod. No. 40-100;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 ACATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 67
 DB 74 ACATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 133
 QY 68 CCAATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 127
 DB 134 CCAATGTCACCAACACGCGGAGTATGCTTTTCATCATCTCGGTCGATATGCGCA 193

QY 128 TGGTAATGCTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 187
 DB 144 TGGTAATGCTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 253
 QY 188 TGGTAATGCTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 247
 DB 254 TGGTAATGCTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 313
 QY 248 ATGATGACAAATCTGCTATAGCAATTTCTGAGTGGGAGCTGAGTGGCTGG 307
 DB 314 ATGATGACAAATCTGCTATAGCAATTTCTGAGTGGGAGCTGAGTGGCTGG 373
 QY 308 CAGCGGCTTTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 367
 DB 374 CAGCGGCTTTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 433
 QY 468 CCGGCTATTTGTAG 482
 DB 434 CCGGCTATTTGTAG 448
 RESULT 14
 AAK25499
 ID AAK25499 standard, DNA, 448 BP.
 AC
 AAK25499;
 06-NOV-2001 (first entry)
 Human: brain expressed single exon probe SEQ ID NO: 54.
 Human: brain expressed single exon probe; probe;
 microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 epilepsy; cancer; ss.
 Homo sapiens.
 W0200157275-A2.
 09-AUG-2001.
 30-JAN-2001: 2001W0-US00667.
 04-FEB-2000: 2000US-020412.
 26-MAY-2000: 2000US-0207456.
 30-JUN-2000: 2000US-0608408.
 03-AUG-2000: 2000US-0632366.
 21-SEP-2000: 2000US-0234697.
 27-SEP-2000: 2000US-0234659.
 04-OCT-2000: 2000US-0024263.
 (M I E) MOLECULAR DYNAMICS INC.
 Penn SD, Hanzel BK, Chen W, Rank DR;
 WPI, 2001: 484446/52.
 Single exon nucleic acid probes for analyzing gene expression in human
 brains -
 Example 4: SEQ ID NO: 54; 458pp + Sequence Listing English
 The present invention provides a number of single exon nucleic acid
 probes which are derived from genomic sequences expressed in the human
 brain. They can be used to measure gene expression in brain cell samples,
 which may enable the diagnosis and improved treatment of nervous system
 diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 epilepsy and cancers. The present sequence is one of the probes of the
 invention

Sequence 448 BP: 79 A: 147 C: 112 G: 110 T: 6 other:
 Query Match 59.2% (over 276) on 22 (expt 419)

Best Local Similarity 100.0%; Pred. No. 40-100;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 ACAATGCAACAAACAGTGGGAGATGCTTTGGTCTTCACTATGCTGCTCA 67
 DB 74 ACAATGCAACAAACAGTGGGAGATGCTTTGGTCTTCACTATGCTGCTCA 133
 QY 68 CCAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 127
 DB 134 CCAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 193
 QY 128 TGGTAATGCTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 187
 DB 194 TGGTAATGCTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 253
 QY 188 TGGTAATGCTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 247
 DB 254 TGGTAATGCTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 313
 QY 248 ATGATGACAAATCTGCTATAGCAATTTCTGAGTGGGAGCTGAGTGGCTGG 307
 DB 314 ATGATGACAAATCTGCTATAGCAATTTCTGAGTGGGAGCTGAGTGGCTGG 373
 QY 308 CAGCGGCTTTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 367
 DB 374 CAGCGGCTTTGTCATGCTGAGTTCATGACAAAGTCAATGCTACTGCTCA 433
 QY 468 CCGGCTATTTGTAG 482
 DB 434 CCGGCTATTTGTAG 448
 RESULT 15
 AAK25499
 ID AAK25499 standard, DNA, 448 BP.
 AC
 AAK25499;
 06-NOV-2001 (first entry)
 Human: bone marrow expressed single exon probe SEQ ID NO: 56.
 Human: bone marrow expressed single exon probe; probe;
 microarray; cancer; leukemia; lymphoma; myeloma; ss.
 Homo sapiens.
 W0200157276-A2.
 09-AUG-2001.
 30-JAN-2001: 2001W0-US00668.
 04-FEB-2000: 2000US-0180312.
 26-MAY-2000: 2000US-0207456.
 30-JUN-2000: 2000US-0608408.
 03-AUG-2000: 2000US-0632366.
 21-SEP-2000: 2000US-0234697.
 27-SEP-2000: 2000US-0234659.
 04-OCT-2000: 2000US-0024263.
 (M I E) MOLECULAR DYNAMICS INC.
 Penn SD, Hanzel BK, Chen W, Rank DR;
 WPI, 2001: 484446/53.
 Human genome-derived single exon nucleic acid probes useful for
 analyzing gene expression in human bone marrow -
 Example 4: SEQ ID NO: 56; 458pp + Sequence Listing English.
 The present invention provides a number of single exon nucleic acid

Tue Dec 31 14:03:11 2002

us-09-727-770-1.rng

cc probes which are derived from genomic sequences expressed in the human
 cc bone marrow. They can be used to measure gene expression in bone marrow
 cc samples, which may enable the improved diagnosis and treatment of cancers
 cc such as lymphoma, leukemia and myeloma. The present sequence is one of
 cc the probes of the invention.

XX
 SQ Sequence 448 bp; 79 A; 147 C; 112 G; 110 T; 0 other;

Query Match	59.2%	Score 375;	DB 22;	Length 448;
Best Local Similarity	100.0%	Pred. No. 4c 100;		
Matches 375;	Conservative 0;	Mismatches 0;	Indels 0;	Caps 0;

QY	8	ATATGTCACAAACAGCCGAGCTATGCTTTGTTTCACACATCTCGGGTGTATGCTCA	67
DB	74	ACATGTCACAAACAGCCGAGCTATGCTTTGTTTCACACATCTCGGGTGTATGCTCA	143
QY	68	CCATGTCCTCCAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	127
DB	134	CCATGTCCTCCAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	193
QY	128	TGGTCATGCTCTGTCATGTGGCCAGAGTGTATGACATGAAGTCTCATCATCGACGTGATCA	187
DB	194	TGGTCATGCTCTGTCATGTGGCCAGAGTGTATGACATGAAGTCTCATCATCGACGTGATCA	253
QY	188	TGGCTGTATCATCACCATCTATGCGCTAGTGGAGGCTGTCTCGGCTGGCAACTGCGCTGA	247
DB	254	TGGCTGTATCATCACCATCTATGCGCTAGTGGAGGCTGTCTCGGCTGGCAACTGCGCTGA	313
QY	248	ATGATGACAAACAGTCTCTATAGTAGTTCCTTCACGCTGAGGAGGTGGGCTGAGTGGGCTGG	307
DB	314	ATGATGACAAACAGTCTCTATAGTAGTTCCTTCACGCTGAGGAGGTGGGCTGAGTGGGCTGG	373
QY	408	CAGGAGGCTTTGGCATCTGTCATCTGAGGACACTGGGAAGTGTGCTGACTGCTGCGACGAGG	367
DB	474	CAGGAGGCTTTGGCATCTGTCATCTGAGGACACTGGGAAGTGTGCTGACTGCTGCGACGAGG	433
QY	368	CCGACTATTGTGAG	382
DB	434	CCGACTATTGTGAG	448

Search completed: December 28, 2002, 09:49:05
 Job time : 266 secs

Copyright (c) 1993 - 2002 Computer M.D.
GenBank version 5.1.1

EM nucleic acid nucleic search, using sw model

Run on: December 28, 2002, 09:22:31, Search time 1982 seconds
(without alignments)
5169.815 Million cell updates/sec

Title: US-09-727-770-1
Perfect score: 633
Sequence: 1 tcaagagacatgcacacaa.....ctgagacatctgaactac 633

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743476 residues

Total number of hits satisfying chosen parameters: 23090992

Minimum DB seq length: 0

Maximum hit seq length: 23090992

Post-processing: Minimum Match 0
Maximum Match 100
Listing first 45 summaries

Database: EST

- 1: em_estba:*
- 2: em_estba:*
- 3: em_estba:*
- 4: em_estba:*
- 5: em_estba:*
- 6: em_estba:*
- 7: em_estba:*
- 8: em_estba:*
- 9: em_estba:*
- 10: em_estba:*
- 11: em_estba:*
- 12: em_estba:*
- 13: em_estba:*
- 14: em_estba:*
- 15: em_estba:*
- 16: em_estba:*
- 17: em_estba:*
- 18: em_estba:*
- 19: em_estba:*
- 20: em_estba:*
- 21: em_estba:*
- 22: em_estba:*
- 23: em_estba:*
- 24: em_estba:*
- 25: em_estba:*
- 26: em_estba:*
- 27: em_estba:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	475	59.2	908	14	BQ722145
2	375	59.2	981	14	BQ678034
3	375	59.2	1091	14	BQ232887
4	373.4	59.0	993	14	BQ463421
5	373.4	59.0	1063	14	BQ053546
6	371.8	58.7	922	14	BQ421511

7	371.8	58.7	1076	14	BQ229295
8	369.6	58.4	986	14	BQ23102
9	369.6	58.4	991	14	BQ918449
10	368.6	58.2	1076	13	BQ544498
11	368.2	58.2	1222	13	BQ478359
12	363	57.3	696	14	BQ791193
13	362.9	57.3	928	14	BQ270185
14	362.4	57.3	1061	13	BQ562656
15	361.8	57.2	1031	13	BQ554691
16	360.2	56.9	996	14	BQ054594
17	360	56.9	1093	13	BQ547917
18	359.8	56.8	991	9	A1524214
19	359.4	56.8	908	14	BQ224964
20	358.8	56.7	1201	14	BQ424665
21	357.8	56.7	872	14	BQ722238
22	357.5	56.5	883	14	BQ713173
23	353	55.8	631	14	BQ769410
24	353	55.8	634	14	BQ843350
25	353	55.8	638	14	BQ818948
26	353	55.8	657	14	BQ769872
27	349.6	55.2	995	14	BQ921094
28	348.6	55.1	1992	13	BQ452232
29	348.2	55.0	1136	13	BQ554912
30	347.6	54.9	841	12	BQ637438
31	347.6	54.9	913	14	BQ949611
32	347.4	54.9	907	14	BQ879221
33	346.6	54.8	722	12	BQ797916
34	345.8	54.6	1017	14	BQ549877
35	345.6	54.6	821	9	A1552541
36	345.6	54.6	969	14	BQ060051
37	345.6	54.6	1169	14	BQ928304
38	344.8	54.5	921	9	A1552543
39	344.4	54.4	670	12	BQ744415
40	343.4	54.2	936	14	BQ919999
41	343.2	54.2	904	14	BQ218416
42	342	54.0	682	12	BQ831575
43	341.2	53.9	1104	14	BQ924578
44	338.2	53.4	645	10	AW514320
45	336.4	53.1	869	9	AL575876

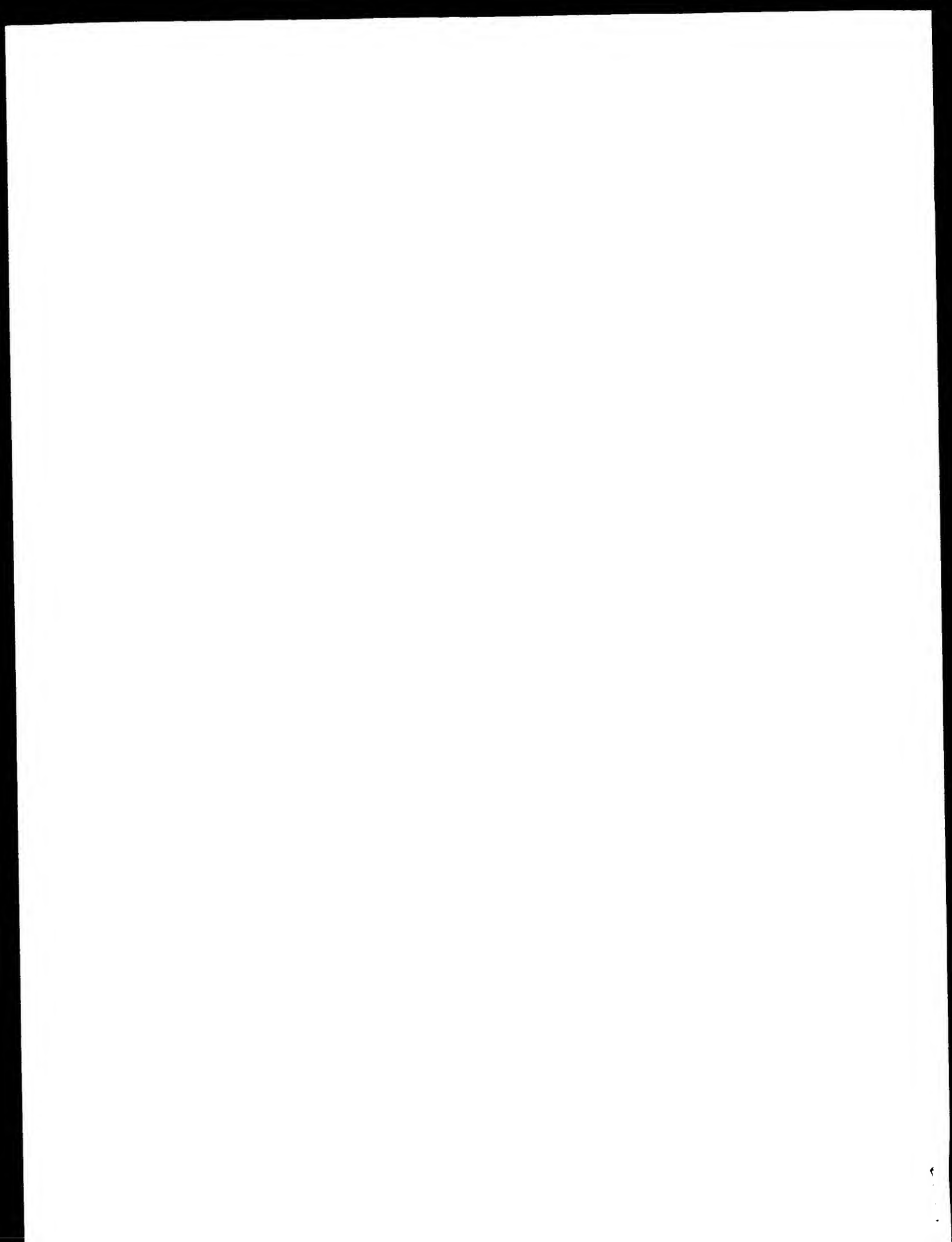
ALIGNMENTS

RESULT 1	BQ722145	908 bp	mRNA	linear	EST 16 JUN 2002
LOCUS	BQ722145	AGRCOURT_8294578	Lupski_sympathetic_trunk	Homo sapiens cDNA clone	
DEFINITION	IMAG:6194244 5', mRNA sequence.				
ACCESSION	BQ722145				
VERSION	BQ722145.1	01:21861042			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIB-MDC help:zmap.cgi?db=genbank
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Stransberg, Ph.D.
Email: rstransb@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LAM13599 row: c column: 13
High quality sequence stop: 644.
Location/Qualifiers
1: 908

Human
SOURCE ORGANISM
Bono sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 886)
Accession: U01001
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rstraus@nsl.nih.gov
Tissue: Prostate; Cell Line: PC-9
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The M.A.G.E. Consortium (LNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the M.A.G.E. Consortium (LNL) at:
http://image.llnl.gov
Plate: LAM3402 row: m column: 17
High quality sequence step: 576,
High quality sequence step: 576,
Location/Qualifiers
1..886
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:605048"
/clone_lib="NH_MGC_72"
/tissue_type="prostate melanoma"
/lab_host="pH10B (phage-resistant)"
/note="origin: skin; Vector: pMW-Script6; Site: 1; Nucle-
Site-2; Salt; Cloned unidirectionally. Primer: 5' end of
Average insert size: 2 kb. Library constructed by Life
Technologies."
BASE COUNT 151 a 314 c 247 g 104 t
ORIGIN
Query Match 58.4%; Score 369.6; DB 14; Length 886;
Best local Similarity 78.3%; Pred. No. 1.6e-85;
Matches 512; Conservative 0; Mismatches 99; Indels 4; Gaps 4;
QY 12 GTCCAAACACACCGCCGAGTATCTTGGTTTCACGATCTCGGGTGGTATGGTACCAT 71
DB 121 GTCCAAAGAGGAGGAGTATCTTGGTTTCACGATCTCGGGTGGTATGGTACCAT 180
QY 72 GCTCTCAGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 131
DB 181 GCTCTCAGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
QY 142 CATCTCTCATGTGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATG 191
DB 241 CATCTCTCATGTGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATG 297
QY 192 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 251
DB 298 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 357
QY 252 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299
DB 358 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 417
QY 390 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 359
DB 418 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 477
QY 419 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 414
DB 478 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 537
QY 415 -----CTCATCTCTCATGATGATGATGATGATGATGATGATGATGATGATG 452
DB 548 CTCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 597
QY 453 CAGTCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 512
DB 598 CAGTCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 657
QY 513 CAGTCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 572
DB 658 ATAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716
QY 573 TCACT 626
DB 717 CCACT 770
FEATURES
SOURCE
1..886
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:605048"
/clone_lib="NH_MGC_72"
/tissue_type="prostate melanoma"
/lab_host="pH10B (phage-resistant)"
/note="origin: skin; Vector: pMW-Script6; Site: 1; Nucle-
Site-2; Salt; Cloned unidirectionally. Primer: 5' end of
Average insert size: 2 kb. Library constructed by Life
Technologies."

High quality sequence step: 715,
Location/Qualifiers
1..1076
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:605048"
/clone_lib="NH_MGC_66"
/tissue_type="large cell carcinoma"
/lab_host="pH10B (phage-resistant)"
/note="origin: lung; Vector: pMW-Script6; Site: 1; Nucle-
Site-2; Salt; Cloned unidirectionally. Primer: 5' end of
Average insert size: 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 169 a 391 c 218 t 1 others
Query Match 58.7%; Score 371.8; DB 14; Length 1076;
Best local Similarity 78.4%; Pred. No. 4.6e-86;
Matches 513; Conservative 0; Mismatches 98; Indels 4; Gaps 4;
QY 12 GTCCAAACACACCGCCGAGTATCTTGGTTTCACGATCTCGGGTGGTATGGTACCAT 71
DB 121 GTCCAAAGAGGAGGAGTATCTTGGTTTCACGATCTCGGGTGGTATGGTACCAT 180
QY 72 GCTCTCAGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 131
DB 181 GCTCTCAGTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
QY 142 CATCTCTCATGTGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATG 191
DB 241 CATCTCTCATGTGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATG 297
QY 192 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 251
DB 298 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 357
QY 252 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299
DB 358 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 417
QY 390 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 359
DB 418 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 477
QY 419 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 414
DB 478 TGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 537
QY 415 -----CTCATCTCTCATGATGATGATGATGATGATGATGATGATGATGATG 452
DB 548 CTCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 597
QY 453 CAGTCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 512
DB 598 CAGTCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 657
QY 513 CAGTCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 572
DB 658 ATAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716
QY 573 TCACT 626
DB 717 CCACT 770
RESULT 8
LOCUS BQ23102 886 bp mRNA linear F01 02-MAY-2002
DEFINITION BQ23102_2592773 NH_MGC_72 Homo sapiens cDNA clone IMAGE:6050488
5' mRNA sequence.
ACCESSION BQ23102
VERSION BQ23102.1 GI:20414502
KEYWORDS EST.



ADDRESSEE: CARELLA, BYRNEL HAIN, GILLILLAN,
ADDRESSEE: CE'CHI, STEWART & OLSTEIN
STREET: 6, BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05363A
FILING DATE: SUBMITTED HERewith
CLASSIFICATION:
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/BOOK NUMBER: 425800-118
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CUNA
ACT-US94-05363A-1

[illegible]

```

RESULT 7
US-09-408-647A-1
: Sequence 1, Application US/0408647A
: Patent No. 6490858
: GENERAL INFORMATION:
: APPLICANT: Kobayashi, Donald
: TITLE OF INVENTION: Chitinase Gene From Streptococcus
: TITLE OF INVENTION: maltophilia
: FILE REFERENCE: Rot-Cook 98-0090
: CURRENT APPLICATION NUMBER: US/09/408,647A
: CURRENT FILING DATE: 1999-08-26
: PRIOR APPLICATION NUMBER: 60/098,046
: PRIOR FILING DATE: 1998-08-27

```

```

1 NUMBER OF SEQ ID NOS: 2
2
3 SOFTWARE: FastSeq for Windows Version 4.0
4
5 SEQ ID NO 1
6
7 LENGTH: 2810
8
9 TYPE: DNA
10
11 ORGANISM: Stenotrophomonas maltophilia
12
13 FEATURE:
14
15 NAME/KEY: CDS
16
17 LOCATION: (306)...(2405)
18
19 OTHER INFORMATION: Open reading frame 1 (orf1)
20
21 NAME/KEY: seq_getid
22
23 LOCATION: (306)...(248)
24
25 NAME/KEY: terminator
26
27 LOCATION: (2457)...(2481)
28
29 NAME/KEY: PBS
30
31 LOCATION: (295)...(298)
32
33 NAME/KEY: CDS
34
35 LOCATION: (314)...(2162)
36
37 OTHER INFORMATION: Open reading frame 2 (orf2)
38
39 PUBLICATION INFORMATION:
40
41 DATABASE ACCESSION NUMBER: Genbank No. 6399958 AF014950
42
43 DATABASE ENTRY DATE: 1997-09-24
44
45 OS-09-2408-647A-1

```

[illegible]

```

1  RESULT 8
2  US 08-698-551-11
3  : Sequence 11, Application US/08698551
4  : Patent No. 5712481
5  : GENERAL INFORMATION:
6  : APPLICANT: Lin, Hui-fang
7  : APPLICANT: Chen, Jennifer H.
8  : APPLICANT: Schievella, Andrea
9  : APPLICANT: Graham, James
10 : TITLE OF INVENTION: NOVEL INF RECEPTOR LEADIN LIGAND
11 : TITLE OF INVENTION: PROTEINS AND INHIBITORS OF LIGAND BINDING
12 : NUMBER OF SEQUENCES: 18
13 : CORRESPONDENCE ADDRESS:
14 : ADDRESSEE: Genetics Institute, Inc.
15 : STREET: 87 Cambridgepark Drive
16 : CITY: Cambridge
17 : STATE: Massachusetts
18 : COUNTRY: USA
19 : ZIP: 02140
20 : COMPUTER READABLE FORM:
21 : MEDIUM TYPE: Floppy disk
22 : COMPUTER: IBM PC compatible
23 : OPERATING SYSTEM: PC-DOS/MS-DOS
24 : SOFTWARE: Patent In Release #1.0, Version #1.25

```


51	16	7.8	1007	8	SCAFBASE	XL15155 S.cerevisiae	124	11	5.4	829	3	AK000947	AK000947 Caco2 bab
52	16	7.8	62643	8	SCF3471	U18630 Saccharomy	125	11	5.4	889	8	pc048465	U18630 Saccharomy
53	13	6.3	159744	2	AC020189	Drosophila	126	11	5.4	918	8	S40059	U18630 Saccharomy
54	13	6.3	196248	3	AC007147	Drosophila	127	11	5.4	1378	8	PC081519	U18630 Saccharomy
55	13	6.3	260791	3	AE004629	AE004629 Drosophila	128	11	5.4	110900	9	HS1392M17	AI049844 Human DNA
56	13	6.3	265574	2	AC087125	Mus muscu	129	10	4.9	1323	8	AR069689	ASPC1411
57	12	5.9	202	8	MZORFED	M95063 Zea mays pu	130	10	4.9	1548	8	SI0473	X55947 S.pombe spo
58	12	5.9	300	11	G71486	G71486 A60078534FM	131	10	4.9	12755	8	AI216846	AF216846 S.pombe chr
59	12	5.9	495	8	AY090271	AY090271 Arabidops	132	10	4.9	42149	8	SPAC1B3	Z08598 S.pombe chr
60	12	5.9	495	8	AY127950	AY127950 Arabidops							
61	12	5.9	501	8	AY098985	AY098985 Arabidops							
62	12	5.9	665	8	ATHAVAPD	U44585 Arabidops							
63	12	5.9	667	8	AY093249	AY093249 Arabidops							
64	12	5.9	680	8	AF022925	AF022925 Vigna rad							
65	12	5.9	689	8	ATHAVAPC	U44584 Arabidops							
66	12	5.9	692	8	ATHAVAPA	U44582 Arabidops							
67	12	5.9	722	8	AB001941	AB001941 Arabidops							
68	12	5.9	749	8	AF424574	AF424574 Arabidops							
69	12	5.9	752	8	AB001940	AB001940 Arabidops							
70	12	5.9	805	6	AX461537	AX461537 Sequence							
71	12	5.9	809	8	CH014669	U14669 Gossypium h							
72	12	5.9	813	8	AB024275	AB024275 Citrus un							
73	12	5.9	820	8	K0816244	U16244 Kalanchoe d							
74	12	5.9	821	8	AS1ATPASHH	M73232 Avena sativ							
75	12	5.9	825	8	AF004938	AF004938 Arabidops							
76	12	5.9	825	8	ATHAVAP	U44581 Arabidops							
77	12	5.9	826	8	AB003942	AB003942 Arabidops							
78	12	5.9	831	8	AF372872	AF372872 Arabidops							
79	12	5.9	833	8	AF416606	AF416606 Penicetu							
80	12	5.9	842	8	AB024274	AB024274 Citrus un							
81	12	5.9	876	8	AF286464	AF286464 Porterosi							
82	12	5.9	879	8	AF412097	AF412097 Arabidops							
83	12	5.9	895	8	GH013670	U13670 Gossypium h							
84	12	5.9	896	8	AF194814	AF194814 Dendrob							
85	12	5.9	910	8	AF331709	AF331709 Aviceenia							
86	12	5.9	911	8	AY086410	AY086410 Arabidops							
87	12	5.9	911	8	AY087112	AY087112 Arabidops							
88	12	5.9	912	8	HVS3VATAP	AY002061 Heta vulg							
89	12	5.9	914	8	AY049249	AY049249 Arabidops							
90	12	5.9	915	8	OSD27098	X95751 N.tabacum m							
91	12	5.9	921	8	NTS0VATAP	AF310228 Lycopersi							
92	12	5.9	946	8	AF010228	AF022925 Vigna rad							
93	12	5.9	956	8	AF022925	X95752 N.tabacum m							
94	12	5.9	980	8	NTPE01A1P	X94999 M.crystalli							
95	12	5.9	982	8	AB024276	AB024276 Citrus un							
96	12	5.9	991	8	MCVATAPC	AY059836 Arabidops							
97	12	5.9	1028	8	AY059836	AB036922 Citrus un							
98	12	5.9	1060	8	AB036922	X98851 H.vulgaris							
99	12	5.9	1097	8	HVS0RCATAP	AB036924 Citrus un							
100	12	5.9	1369	8	AB036924	AX461536 Sequence							
101	12	5.9	1524	6	AX461536	AB036923 Citrus un							
102	12	5.9	1909	8	AB046923	U44583 Arabidops							
103	12	5.9	2546	8	ATHAVAPB	AF153677 Arabidops							
104	12	5.9	2762	8	AF153677	AC007195 Arabidops							
105	12	5.9	3918	8	AT059757	AC120526 Oryza sat							
106	12	5.9	56284	8	AC007195	AL035679 Arabidops							
107	12	5.9	96466	2	AC120526	AF153677 Arabidops							
108	12	5.9	100466	2	AF153677	AC104285 Oryza sat							
109	12	5.9	110913	2	CH037502	AC104285 Oryza sat							
110	12	5.9	117541	2	AC104285	AL023094 Arabidops							
111	12	5.9	11942	8	AC007797	AC006444 Genomic s							
112	12	5.9	125502	8	AT14120	AC092916 Homo sapi							
113	12	5.9	129547	8	AC096434	AC095460 Oryza sat							
114	12	5.9	132182	9	AC092916	AF004775 Oryza sat							
115	12	5.9	135689	2	AC095460	AC129276 Partis no							
116	12	5.9	137877	2	AP004775	AC095460 Oryza sat							
117	12	5.9	146571	2	AC128276	AL161586 Arabidops							
118	12	5.9	190721	8	AP003523	AC026009 Homo sapi							
119	12	5.9	195155	8	AT3HE18A2	AL161594 Arabidops							
120	12	5.9	196640	2	AC026009	AL128647 Oryza sat							
121	12	5.9	198987	8	AT3HE18A2	AX235195 Sequence							
122	12	5.9	244803	2	AC128647								
123	11	5.4	45	6	AX235195								

ALIGNMENTS

RESULT 1	AX337957	Sequence 1	From patent WO/94 380.	LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
DEFINITION	AX337957	Sequence 1	From patent WO/94 380.	LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
ACCESSION	AX337957	Sequence 1	From patent WO/94 380.	LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
VERSION	AX337957.1	GI:18128674		LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
KEYWORDS	human.			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
ORGANISM	Homo Sapiens			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
SOURCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
REFERENCE	1	Li, Z., Wei, M.H., Ketchum, K., and Beasley, E.M.		LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
AUTHORS	Isolated human transporter proteins, nucleic acid molecules encoding human transporter proteins, and uses thereof			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
TITLE	Patent: WO 0194480-A 1 13-DEC-2001			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
JOURNAL	Applet Corporation Robert A. Millman Assistant Secretary (US)			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
FEATURES	location/Qualifiers			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
source	1..633			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
BASE COUNT	128 a 202 c 153 q 150 t			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
ORIGIN				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
Alignment Scores:				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
Pred. No.:	1936-192			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
Score:	205.00			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
Percent Similarity:	100.00%			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
Best Local Similarity:	100.00%			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
Query Match:	6			LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
DB:				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
US-09-727-770 2 (1-205) x AX337957 (1-633)				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
QY 1 MetSerAsnAsnSerProGluTyrAlaGluValAlaThrThrLeuSerGlyAlaMetAlaThr 20				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
DB 10 ATCTCCAC 69				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
QY 21 MetValSerSerGlyLeuGlyAlaAlaGlySerGlyMetAlaGlySerGlyThrGlyLeuMet 40				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
DB 70 ATGCTCTCCAC 129				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
QY 41 AlaMetSerValMetTrpProGluGluGluMetMetLysSerThrLeuValAlaThrMet 60				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
DB 130 GCCATGCT 189				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
QY 61 AlaGlyLeuLeuThrLeuTyrGlyLeuValAlaAlaGluGluGluGluGluGluGluGlu 80				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
DB 190 GCCTGATTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 249				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
QY 81 AsnAsnAsnSerLeuTyrSerSerProGluGluGluGluGluGluGluGluGluGluGlu 100				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
DB 250 GATGAC 409				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
QY 101 AlaGlyPheAlaThrLeuValLeuValGlyAspGluGluGluGluGluGluGluGluGlu 120				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
DB 310 GCCGCT 469				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002
QY 121 ArgLeuPheValGlyMetLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140				LOCUS	AX337957	633 bp	DNA	Human	PAT 09 JAN 2002

Tue Dec 31 14:03:13 2002

us-09-727-770-2.olil0p2n.rge

Peterson, K., Plunkkarn, P., Picot, N., Pollard, V., Raymond, C.,
 Rott, R., Rieback, M., Kelly, R., Rice, C., Roy, P., Roman, J.,
 Rossetti, K., Roy, A., Santos, E., Schauer, S., Schupbach, R., Seaman, S.,
 Soverly, P., Spencer, P., Stange-Thomann, N., Stellanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testayo, S., Theodore, J.,
 Topfman, K., Truett, G., Travis, N., Trullio, J., Vassiliou, H.,
 Viel, R., Vo, A., Wilson, R., Wu, X., Wyman, D., Ye, W., Young, G.,
 Zainoum, J., Zembek, J., Zimmer, A., and Zody, M.
 Direct Submission
 Submitted (07-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 on Mar 7, 2002 this sequence version replaced 41:978674.
 All repeats were identified using RepeatMasker.
 Smith, A.F.A. & Green, P. (1996, 1997)
 http://ftp.genetics.wisc.edu/genbank/ftp/ncbi/Marker.htm

Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WMR
 Web site: http://www.scri.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information

Center project name: L485
 Center clone name: 177_h5
 ----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.60741
 Consensus quality: 14221 bases at least Q40
 Consensus quality: 14562 bases at least Q30
 Consensus quality: 147521 bases at least Q20
 Insert size: 13800; agarose-gel
 Insert size: 150099; sum-of-continues
 Quality coverage: 6.5 in Q20 bases; adarose-gel
 Quality coverage: 6.0 in Q20 bases; sum-of-continues

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 20 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 41140: contig of 41140 bp in length
 * 41131 41240: gap of 100 bp
 * 41231 42274: contig of 1044 bp in length
 * 42275 42374: gap of 100 bp
 * 42375 44180: contig of 1805 bp in length
 * 44181 44280: gap of 100 bp
 * 44281 45329: contig of 1049 bp in length
 * 45330 45429: gap of 100 bp
 * 45430 47605: contig of 2176 bp in length
 * 47606 47705: gap of 100 bp
 * 47706 50517: contig of 2812 bp in length
 * 50518 50617: gap of 100 bp
 * 50618 54448: contig of 3831 bp in length
 * 54449 54548: gap of 100 bp
 * 54549 58969: contig of 4421 bp in length
 * 58970 59069: gap of 100 bp
 * 59070 63507: contig of 4438 bp in length
 * 63508 63607: gap of 100 bp
 * 63608 66665: contig of 3058 bp in length
 * 66666 66765: gap of 100 bp
 * 66766 71248: contig of 4483 bp in length
 * 71249 71348: gap of 100 bp
 * 71349 76193: contig of 4845 bp in length
 * 76194 76293: gap of 100 bp
 * 76294 86300: contig of 10007 bp in length
 * 86301 86400: gap of 100 bp
 * 86401 95039: contig of 8639 bp in length
 * 95040 95139: gap of 100 bp
 * 95140 102689: contig of 7550 bp in length

* 102690 102789: gap of 100 bp
 * 102790 113801: contig of 11012 bp in length
 * 113802 113901: gap of 100 bp
 * 113902 122360: contig of 8459 bp in length
 * 122361 122460: gap of 100 bp
 * 122461 132544: contig of 10084 bp in length
 * 132545 132644: gap of 100 bp
 * 132645 147910: contig of 15266 bp in length
 * 147911 148010: gap of 100 bp
 * 148011 151999: contig of 4989 bp in length.
 FEATURES
 Location/Qualifiers
 1..151999
 /contigname="us-09-727-770-2.olil0p2n.rge"
 /db_xref="taxon:9606"
 /label="RP11:1770p"
 /clone="1770p" /accession="us-09-727-770-2.olil0p2n.rge"
 1..41140
 /note="assembly_fragment"
 /clone_end:896
 /vector_side:left
 41231..42274
 /note="assembly_fragment"
 42375..44180
 /note="assembly_fragment"
 44281..45329
 /note="assembly_fragment"
 45430..47605
 /note="assembly_fragment"
 47706..50517
 /note="assembly_fragment"
 50618..54448
 /note="assembly_fragment"
 54549..58969
 /note="assembly_fragment"
 59070..63507
 /note="assembly_fragment"
 63608..66665
 /note="assembly_fragment"
 66766..71248
 /note="assembly_fragment"
 71349..76193
 /note="assembly_fragment"
 76294..86300
 /note="assembly_fragment"
 86401..95039
 /note="assembly_fragment"
 95140..102689
 /note="assembly_fragment"
 102790..113801
 /note="assembly_fragment"
 113902..122360
 /note="assembly_fragment"
 122461..132544
 /note="assembly_fragment"
 132645..147910
 /note="assembly_fragment"
 148011..151999
 /note="assembly_fragment"
 /clone_end:17
 /vector_side:right
 46160 1 1908 others
 BASE COUNT 46958 a 48119 c 48854 g 151999
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,720-191 Length: 151999
 Score: 205.00 Matches: 205
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-727-770-2 (1-205) x NC005973 (1-151999)

us-09-727-770-2.olil0p2n.rge

Tue Dec 31 14:03:13 2002

```

* 17898 18949: contig of 1052 bp in length
* 18950 19049: gap of 100 bp
* 19050 20494: contig of 1445 bp in length
* 20495 20594: gap of 100 bp
* 20595 21775: contig of 1181 bp in length
* 21776 21875: gap of 100 bp
* 21876 23459: contig of 1584 bp in length
* 23460 24599: gap of 100 bp
* 24600 25831: contig of 2272 bp in length
* 25832 25931: gap of 100 bp
* 25932 27763: contig of 1832 bp in length
* 27764 27863: gap of 100 bp
* 27864 29866: contig of 2123 bp in length
* 29867 30086: gap of 100 bp
* 30087 32675: contig of 2589 bp in length
* 32676 32775: gap of 100 bp
* 32776 36266: contig of 3491 bp in length
* 36267 36366: gap of 100 bp
* 36367 39687: contig of 3321 bp in length
* 39688 42811: contig of 3024 bp in length
* 42812 42911: gap of 100 bp
* 42912 47748: contig of 4837 bp in length
* 47749 47848: gap of 100 bp
* 47849 52056: contig of 4208 bp in length
* 52057 52156: gap of 100 bp
* 52157 56257: contig of 4101 bp in length
* 56258 56357: gap of 100 bp
* 56358 62535: contig of 6178 bp in length
* 62536 62645: gap of 100 bp
* 62646 65138: contig of 2503 bp in length
* 65139 65238: gap of 100 bp
* 65239 69640: contig of 4402 bp in length
* 69641 69740: gap of 100 bp
* 69741 74838: contig of 5098 bp in length
* 74839 74938: gap of 100 bp
* 74939 79482: contig of 4544 bp in length
* 79483 79582: gap of 100 bp
* 79583 81888: contig of 2106 bp in length
* 81889 81988: gap of 100 bp
* 81889 89859: contig of 7871 bp in length
* 89860 89959: gap of 100 bp
* 89960 158380: contig of 68421 bp in length
* 158381 158480: gap of 100 bp
* 158481 166923: contig of 8443 bp in length
* 166924 167023: gap of 100 bp
* 167024 176604: contig of 9881 bp in length
* 176605 177004: gap of 100 bp
* 177005 184595: contig of 7591 bp in length.
FEATURES
    Location/Qualifiers
        1..184595
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="17"
        /map="17"
        /clone="RP11-81A22"
        /clone_lib="RP11 Human Male BAC"
misc_feature
    1..10663
        /note="assembly_fragment"
        clone_end:586
        vector_side:left
        10764..11392
            /note="assembly_fragment"
            11493..12334
                /note="assembly_fragment"
                12335..13392
                    /note="assembly_fragment"
                    13433..14716
                        /note="assembly_fragment"
                        14817..16101
                            /note="assembly_fragment"
                            16202..17797
                                /note="assembly_fragment"

```

```

17898..18949
/note="assembly_fragment"
19050..20494
/note="assembly_fragment"
20595..21775
/note="assembly_fragment"
21876..23459
/note="assembly_fragment"
23460..24599
/note="assembly_fragment"
24600..25831
/note="assembly_fragment"
25832..27763
/note="assembly_fragment"
27864..29866
/note="assembly_fragment"
30087..32675
/note="assembly_fragment"
32776..36266
/note="assembly_fragment"
36367..39687
/note="assembly_fragment"
39788..42811
/note="assembly_fragment"
42912..47748
/note="assembly_fragment"
47849..52056
/note="assembly_fragment"
52157..56257
/note="assembly_fragment"

Alignment Scores:
Pred. No.: 1,120-190 Length: 184595
Score: 205.00 Matches: 205
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 Indels: 0
DB: 2 Gaps: 0

US-09-727-770-2 (1-205) x AC015913 (1-184595)
QY 1 MetSerAspAsnSerProGluTyrAlaLeuValPheThrIleSerGlyAlaMetAlaThr 20
Db 42060 ATGTCCACACACACACACACACACACACACACACACACACACACACACACACAC 42119
QY 21 MetValSerSerGlyLeuGlyAlaAlaCysGlyMetAlaTyrAsnGlyThrGlyThrMet 40
Db 42120 ATGGTCCTCACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 42179
QY 41 AlaMetSerValMetTrpProGluLeuIleHisMetLysSerIleIleProValValMet 60
Db 42180 GGCATGTCCTCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 42239
QY 61 AlaGlyIleIleThrIleTyrGlyLeuValAlaAlaValProProAlaAsnSerLeuAsn 80
Db 42240 GGTGGTATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 42299
QY 81 AspAspAsnSerLeuTyrSerSerPheLeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
Db 42300 GATGACACACAGTCTCATATACACATTCCTCCACCTGACCTGACCTGACCTGACCT 42359
QY 101 AlaGlyPheAlaIleValIleValGlyAspThrGlyLysCysGlyThrAlaGluIlePro 120
Db 42360 GCGGGCTTTCGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACG 42419
QY 121 AratLeuPheValGlyMetIleLeuIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140
Db 42420 CGAATTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 42479
QY 141 LysGluProLeuSerLysProThrSerHisSerIleAlaCysLysAspHisProSerSer 160
Db 42480 AAGCAGCGCCCTCTCAAAACCCACGACGACGACGACGACGACGACGACGACGACGACG 42539
QY 161 PheArgAsnLysGluProAspThrIleValIleGlyLysThrIleProSerValValAspLeu 180
Db 42540 TTCGCGCAACAACACACACACACACACACACACACACACACACACACACACACACAC 42599

```


QUERY MATCH: 9 278 Indels: 0
 DB: 8 Gaps: 0

US-09-727-770-2 (1-205) x ENSO10N1 (1-636)
 QY 115 GlythralaciladInProArqleupheValGlyMetIleLeuIleLeuIleLeuAla 133
 DB 471 GGAACAGCACAAACACCGCTTTCGCGAATGATCTTATTTCATTTTCGCC 527

RESULT 11
 CRG10CKN
 LOCUS Botrytis cinerea strain #4 cDNA library under conditions of
 DEFINITION nitrogen deprivation.
 ACCESSION All15647.1 GI:5830863
 VERSION All15647.1 GI:5830863
 KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botrytis cinerea fockeliana
 ORGANISM Botrytis cinerea fockeliana
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
 Helotiales; Sclerotiniaceae; Botryotinia.
 REFERENCE 1 (bases 1 to 696)
 AUTHORS Bliton, F., Levis, C., Fattini, B., Fradier, J.M. and Puyguez, Y
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
 78026 Versailles, France
 REFERENCE 2 (bases 1 to 696)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage ;
 CP 5706 91057 Evry cedex FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES
 source
 1..696
 /organism="Botrytis cinerea fockeliana"
 /strain="74"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W32A071"

BASE COUNT 155 a 166 c 145 g 230 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,530-09 Length: 656
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.27% Indels: 0
 DB: 8 Gaps: 0

US-09-727-770-2 (1-205) x ENSO10N1 (1-636)
 QY 115 GlythralaciladInProArqleupheValGlyMetIleLeuIleLeuIleLeuAla 133
 DB 468 GGAACAGCACAAACACCGCTTTCGCGAATGATCTTATTTCATTTTCGCC 524

RESULT 12
 AF008924
 LOCUS Aedes aegypti V-ATPase C-subunit mRNA complete cds.
 DEFINITION AF008924
 ACCESSION AF008924.1 GI:2454489
 VERSION AF008924.1 GI:2454489
 KEYWORDS Aedes aegypti.
 SOURCE Aedes aegypti.
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
 REFERENCE 1 (bases 1 to 854)

GILL, S.S., Chu, P.B., Smethurst, P., Pietrantoni, P.V. and Ross, J.S.
 Isolation of the V-ATPase A- and C-subunit cDNAs from mosquito
 midgut and Malpighian tubules
 Arch. Insect Biochem. Physiol. (1997) In press
 2 (bases 1 to 854)
 GILL, S.S.
 Direct Submission
 Submitted (17-JUN-1997) Entomology, U. California, Riverside, 92519
 Boyce Hall, Riverside, CA 92521, USA
 Location/Qualifiers
 1..854
 /organism="Aedes aegypti"
 /db_xref="taxon:7159"
 163..636
 /note="proteolipid"
 /codon_start=1
 /product="V-ATPase C-subunit"
 /protein_id="AAB71660.1"
 /db_xref="GI:2454489"
 /translation="MALLTLLFVYDTPVVM*AAATLGGALCAAYTAKSTGGLAAM
 SYMRCLLMKSLFVVMAGLIALYGVAVVATAGSLDTPKSTYLPGLIHGAGLAVG
 ISGLAAGFAIGLVGLAGVGTAGQRIFGVMILLLFAEVLGLGLVAVLYLTK*"

BASE COUNT 207 a 248 c 196 g 203 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,096-09 Length: 854
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.27% Indels: 0
 DB: 3 Gaps: 0

US-09-727-770-2 (1-205) x A10637.4 (1-854)
 QY 115 GlythralaciladInProArqleupheValGlyMetIleLeuIleLeuIleLeuAla 133
 DB 526 GGTACTGGCCAGCAGCAACACGCTTTCGCGAATGATCTTATTTCATTTTCGCC 582

RESULT 13
 AY099523
 LOCUS Danio rerio vacuolar ATP synthase 16 kDa proteolipid subunit
 DEFINITION (atp6j) mRNA, complete cds.
 ACCESSION AY099523.1 GI:20977566
 VERSION AY099523.1 GI:20977566
 KEYWORDS zebrafish.
 SOURCE Danio rerio
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 902)
 AUTHORS Gelling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, P.,
 Chen, W., Burgess, S., Haldi, M., Artzt, K., Farnham, S., Liu, S., Y.,
 Nissen, K.M. and Hopkins, N.
 TITLE Essential role for early zebrafish rapidly identifies genes
 JOURNAL Nat. Genet. 31 (2): 135-149 (2002)
 MEDLINE 22038002
 PubMed 12009978
 REFERENCE 2 (bases 1 to 902)
 AUTHORS Amsterdam, A., Gelling, G., and Hopkins, N.
 TITLE Direct Submission
 JOURNAL Submitted (24-APR-2002) Center for Cancer Research, Massachusetts
 Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139,
 USA
 Location/Qualifiers
 1..902
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 1..902
 /gene="atp6j"

US-09-727-770-2 (1-205) x MM013842 (1-1134)	
QY 115	glyThrAlaGlnGlnProArqLeuPheValGlyMetHleLeuHlleLeuHlePheAla 133
DB 480	GGCAATGGGACAGCTGGACTTTTGTGACATGATGCTGATGCTGATGCTGATGCTG 536
RESULT 20	
LOCUS	BMVATPPLC
DEFINITION	bovine proteolipid protein of the H ⁺ -ATPase of chromaffin granules
ACCESSION	J03835 M61709
VERSION	J03835.1 GI:162714
KEYWORDS	proteolipid protein of H ⁺ -ATPase,
SOURCE	bovine (adult) adrenal medulla chromaffin granule membrane, cDNA to mRNA.
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
AUTHORS	Mandel, M., Moriyma, Y., Holmes, J.D., Fan, Y.C., Nelson, H. and Nelson, N.
TITLE	cDNA sequence encoding the 16-kDa proteolipid of chromaffin granules implies gene duplication in the evolution of H ⁺ -ATPases
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 85 (15): 5521-5524 (1988)
MEDLINE	88299753
PubMed	2456571
KEYWORDS	2 (bases 1 to 1133)
AUTHORS	Nelson, N.
JOURNAL	unpublished (1991)
COMMENT	draft entry and printed copy of sequence for [1] kindly provided by N. Nelson, 06/23/88.
FEATURES	Location/Qualities
source	1..1133
mpna	/organism="Bos taurus"
	/db_xref="taxon:9913"
	/c1..1133
CDS	/product="PLP mRNA"
	133..600
	/note="proteolipid protein of H ⁺ -ATPase"
	/codon_start=1
	/db_xref="GI:162715"
	/translation="MSEAKNGPEYAFVMEGASAMVFSALCAAGCTAKSGTGIAAM SVRPEIMKMSITPVVMAGLIAIYGLVAVILANSUNDGISLYRSFLQAGISVGLS GLAAGFAIGVGLAGVGTAGQPRLFVGMILILIFAELVGLGLIVALLITK"
BASE COUNT	183 a 384 c 314 g 252 t
ORIGIN	Unreported.
Alignment Scores:	
Prod. No.:	5,01e-09
Score:	19.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	9.27%
DB:	4
US-09-727-770-2 (1-205) x BMVATPPLC (1-1133)	
QY 115	glyThrAlaGlnGlnProArqLeuPheValGlyMetHleLeuHlleLeuHlePheAla 133
DB 490	GGACACGACAGACAGCGGCTCTTCTGACATGATGCTGATGCTGATGCTGATG 546
RESULT 21	
LOCUS	BMVHATP
DEFINITION	bovine proteolipid protein of the H ⁺ -ATPase
ACCESSION	X55979
VERSION	X55979.1 GI:8811
KEYWORDS	ATPase; H ⁺ -ATPase; vacuolar H ⁺ -ATPase
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Insecta; Diptera; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	St. Rauber, R.
TITLE	Direct Submission
JOURNAL	Submitted (01-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MG), Cancer Genomics Center, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	Non-Bacterial DNA, 1133 bp, 12,000-2000
COMMENT	Contact: MGC help desk
	Email: mgc@nci.nih.gov
	Tissue Procurement: AIC/Cybio/DIP
	cDNA Library Preparation: Rubin Laboratory
US-09-727-770-2 (1-205) x BMVHATP (1-1138)	
QY 115	glyThrAlaGlnGlnProArqLeuPheValGlyMetHleLeuHlleLeuHlePheAla 133
DB 486	GGACACGACAGACAGCGCTCTTCTGACATGATGCTGATGCTGATGCTGATG 542
RESULT 22	
LOCUS	BM007389
DEFINITION	Homo sapiens, ATPase, H ⁺ transporting, lysosomal (vacuolar) proton pump) 16kb, clone MGC:16271 IMAGE:4831016, mRNA, complete cds.
ACCESSION	BM007389
VERSION	BM007389.1 GI:13948443
KEYWORDS	MGC.
SOURCE	Homo sapiens
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	1 (bases 1 to 1143)
TITLE	Direct Submission
JOURNAL	Submitted (01-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MG), Cancer Genomics Center, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	Non-Bacterial DNA, 1143 bp, 12,000-2000
COMMENT	Contact: MGC help desk
	Email: mgc@nci.nih.gov
	Tissue Procurement: AIC/Cybio/DIP
	cDNA Library Preparation: Rubin Laboratory

Chromosome 16, map position p13.4.

```

Alignment Scores:
  Length: 1162
  Score: 19.00
  Percent Similarity: 100.00%
  Percent Local Similarity: 100.00%
  Query Match: 9.27%
  Gaps: 0

US-09-727-770-2 (1-205) x HUMFCHS06-A (1-1162)

QY 115 GlyThrAlaGlnGlnProArgLeuPheValGlyMetIleLeuIleLeuPheAla 133
|||||
DB 588 GCACGGCCAGCAGCCCGCAGTATTCGTGGCGATGATCCGTATTCATCITGCC 644

RESULT 25
CID002877
GENUS Candida
DEFINITION Candida tropicalis, vacuolar ATPase subunit c gene, complete cds.
ACCESSION U02877
VERSION U02877.1 GI:410089
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Gu, B.-H., Xu, J., and Bean, G. F.
TITLE Structure and expression of the vacuolar ATPase subunit c gene from Candida tropicalis.
JOURNAL Eukaryotic Cell
AUTHORS Fuka, J. A., Funk, J., Ascom, J. L., and Bean, G. F.
TITLE Saccharomyces, mitospore, and vacuolar ATPase subunit c gene from Candida tropicalis.
JOURNAL J Biol Chem
AUTHORS Bean, G. F.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-1994) Bean G. F., University of Cincinnati College of Medicine, Molecular Genetics, Biochemistry, and Microbiology, 231 Bethesda Avenue, Cincinnati, OH 45267-0524, USA
FEATURES
source 1..1386
     /organism="Candida tropicalis"
     /db_xref="taxon:5482"
     join(244..257,645..1103)
     /standard_name="V-ATPase subunit c"
     /codon_start=1
     /evidence=experimental
     /product="vacuolar ATPase subunit c"
     /protein_id="AAA03446.1"
     /db_xref="GI:410090"
     /translation="MSDLCHVYAFETFSIGCAAAIVFTGCAASYTAKSCGVGICATLSV
     TRPDLVKNVVPVPMAGI;AIYGLVSVISVLSISQKQALYTGFTLQAGLSVGLSLCL
     AAGFAIGVIGCAVGRTPAQQRKLVGMILILFAEVLIGLYGLIVALLINSRASQGVTC
     "

intron 258..634
BASE COUNT 417 a 210 c 240 q 519 t
ORIGIN
Alignment Scores:
  Length: 1486
  Score: 19.00
  Percent Similarity: 100.00%
  Percent Local Similarity: 100.00%
  Query Match: 9.27%
  Gaps: 0

US-09-727-770-2 (1-205) x CT002877 (1-1486)

QY 115 GlyThrAlaGlnGlnProArgLeuPheValGlyMetIleLeuIleLeuPheAla 133
|||||
DB 972 GGTACTGCTGCACACACACAGATTATTCGTCGGATACGATTCGATCTTGATTTTCGCT 1028

RESULT 26
MSVATP16S
```


Tue Dec 31 14:03:13 2002

Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission
2 (bases 1 to 114529)
Worley, K.C.
Direct Submission
Submitted (24-sep-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 114529)
Worley, K.C.
Direct Submission
Submitted (12-Jul-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced all:18148582.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

us-09-727-770-2 (1-205) x AC004438 (1-83340)

QY 115 G1YThrAlaGlnInProArqleupheValGlyMetIleLeuIleLeuIlePheValA 133

114529 bp DNA *** SEQUENCING IN PROGRESS
AC098526
Rattus norvegicus clone CH230-9A22, 114529 bp DNA *** SEQUENCING IN PROGRESS
***, 63 unordered pieces.
AC098526
AC098526.6 G1:21729576
HTG: HTGS-PHASE1
Norway rat
Rattus norvegicus
Eukaryota: Metazoa, Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Sciurognathi: Muridae: Murinae:
Rattus.

RESULT 46

AC098526

LOCUS

DEFINITION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 114529)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Alsbrooks, S.L., Anaratunga, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbiera, J., Benton, J., Bimaga, K., Blankenship, K., Bonnin, D.,
Bouk, J., Bowie, S., Brivea, M., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burdick, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chace, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathine, S.R., David, P.,
DeLaney, M.L., Davis, C., Davy, Carroll, L., Dederich, D.A.,
DeLaney, K.R., Delgado, O., Dunn, A.L., Ding, Y., Ditt, H.H.,
Douthwaite, K.J., Draper, H., Duan, Roeha, S., Durbin, K.J.,
Earhart, C., Edgar, B., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, B., Flaugh, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gortell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B.E.,
Honsi, F., Howard, S., Huber, J., Huley, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karissom, E., Kelly, S., Khan, U., King, L., Korvath, J., Kovar, C.,
Kratovic, J., Kresh, A., Landry, N., Leal, B., Lewis, C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louisied, H.,
Lokadi, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, P., Mawhinney, E., McLeod, M.P., Meador, M., Mei, C., Metzger, M.,
Miner, G., Minor, Z., Mitchell, T., Mohabbat, K., Morahan, M., Morris, S.,
Mosser, M., Neal, D., Newton, J., Newton, N., Nuygen, A., Nguyen, N.,
Oragunye, N., Oyedero, K., Pace, A., Payton, B., Perry, J., Perez, L.,
Peters, L., Pickens, R., Primm, E., Pu, L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rohubaku, L., Rolfe, M., Ruiz, S., Savary, G.,
Scherrer, S., Scott, G., Shen, H., Shostari, N., Sisson, I.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tan, H.,
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Uemami, K., Vasquez, L., Vera, V., Villalob, B., Villalob, S.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Woodson, S., Worley, K.,

NOTE: Estimated insert size may differ from sequence length
(see http://www.hbase.com/seq/seqs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 63 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1183: contig of 1183 bp in length
1184: gap of unknown length
1284: contig of 1020 bp in length
1284: gap of unknown length
2404: contig of 1459 bp in length
3762: gap of unknown length
3763: gap of unknown length
4932: contig of 1070 bp in length
5032: gap of unknown length
5625: contig of 1534 bp in length
6253: gap of unknown length
8090: contig of 1465 bp in length
8190: gap of unknown length
8191: contig of 1402 bp in length
9592: gap of unknown length
9593: contig of 1522 bp in length
11215: gap of unknown length
11215: contig of 1657 bp in length
12971: gap of unknown length
13071: contig of 1127 bp in length
14168: gap of unknown length
14298: contig of 1248 bp in length
14199: gap of unknown length
15546: gap of unknown length
15547: contig of 1269 bp in length
16915: gap of unknown length
16916: contig of 1087 bp in length
17016: gap of unknown length
18072: contig of 1419 bp in length
18173: gap of unknown length
19491: gap of unknown length
20746: contig of 1155 bp in length
20846: gap of unknown length

[illegible][illegible]

ADAMS, M. D., GUTHRIE, D. L., BEIT, R. A., EMMER, J. A., SOGINO, J. D.,
ANASTASIDES, P. G., SCHERER, S. E., HILL, P. W., BOSKINS, R. M., GALLIE, R. F.,
CHAPMAN, P. A., JONES, S. F., PUGH, R. L., ASHINGTON, M. A., MACHINSON, S. N.,
SULTON, G. G., WORTEN, J. K., VANDER, J. M. D., ZHANG, Q., CHEN, L. X.,
BRANDON, P. C., ROGERS, Y. H., BLAZEL, K. G., CHAMPE, M., PFEIFFER, B. D.,
WANG, Y. M., LEGG, J., FALTER, F. S., HOLT, G., NELSON, C. R., GABOR
MILLON, C. J., ABRIL, J. F., ANKAYANI, A., AN, H. J.,
ANDREWS-ANDREWS, C. E., HEDIN, P., BALLEW, E. M., PASO, A.,
BAWAZ, K. A., J. Bayraktaroglu, L., Beck, J. M., Besson, K. Y.,
BONIS, P. V., BERMAN, B. P., BHANDARI, D., BOLSHAKOV, S., BERKOW, D.,
BOURKE, J. M., ROBERTS, P., BROTHIER, P., BURTIS, K. C.,
BUSAM, D. A., BUTLER, H., CADORE, R., CENTER, A., CHANDRA, I.,
CHARTY, J. M., CHAVEZ, S., PAHLER, C., D'ARCY, B., GABOR, J. D.,
FABES, B., DELCHER, A., GREGG, J., HAYS, A. B., NEW, J., DICKINSON, M.,
FEDER, E., DE, P. L., FENG, W., FINE, E. H., FINE, B. C., FINE, B. C.,
FURN, P., PURIN, K. L., FRANGOLILLAS, C. C., FORTAG, C., FERREIRA, S.,
FRANCAVALLO, W., FUELE, C., GABRIELIAN, A. B., GARD, R. S.,
GILBERT, W. M., GLASSER, K., GLOCK, A., GOLF, F., GUTIEL, J. R., GUNZ,
QUAN, Y., HARRIS, M., HARTIS, N. L., HARVEY, P., HELMAN, J. J.,
HORNANDER, J. P., HOUK, J., HOSTIN, P., HOUSTON, K. A., HOWLAND, T. J.,
WEL, B. H., HEDMAN, J. D., JALALI, M., KALASH, I., KARPEN, S. H., KERR,
KENNEDY, J. A., KETCHEM, F. A., KIMMEL, R. E., KODIRA, S. E., KRAIT, C.,
KLEIN, S. G., KAP, E., LARSEN, L. K., LEE, Y., LELAND, A. A.,
LIT, J., LIZO, Z., LIANG, Y., LIU, X., LIU, X., MATTEI, R., MCINTOSH, J. C.,
MADON, M. J., MATHIAS, M., MATHIAS, S., MATHIAS, R. V., MATHIAS, C.,
MORRIS, J., MOSHREJI, A., MOUNT, S. M., MOY, M., MURPHY, B., MURPHY, L.,
MURPHY, L., NELSON, J. L., NELSON, E. F., NELSON, K. A., NELSON, K.,
NUSSEN, P. E., PACHECO, J. M., PACHECO, J. M., PACHECO, J. M., PACHECO, J. M.,
POLLARD, J. D., PURI, V., RABIN, M. G., RABIN, M. G., REMINGTON, K.,
SANDERS, P. D., SCHWARTZ, P., SKEN, H., SHU, P. C., SIDEN, KIMOS, I.,
STAMPSON, M., SKUPSKI, M. P., SMITH, T., SPIER, E., SPRADLING, A. C.,
VAL, C. M., ZHANG, Y., ZHANG, Y., ZHANG, Y., ZHANG, Y., ZHANG, Y.,
VENTER, E., WANG, A. H., WANG, X., WANG, Z. Y., WASSARMAN, D. A.,
WEISSBERG, J. M., WEISSBERG, J. M., WEISSBERG, J. M., WEISSBERG, J. M.,
WOLFEY, P. J., WU, Y., YANG, S., YANG, S., YANG, S., YANG, S.,
XIAO, J. N., ZHANG, W., ZHANG, X., ZHANG, X., ZHANG, X., ZHANG, X.,
GIBBS, R. A., MYERS, R. W., WATSON, G. M., and VENTER, J. C.,
The genome sequence of *Drosophila melanogaster*
Science, 287 (341), 216-219, (2000)
20196006
10731132
2 (bases 1 to 345026)
ANASTASIDES, P. G., BRADTON, R. C., ROGERS, Y. H., AND, H. J., BALDWIN, D.,
BRANDON, J., BESSON, K. Y., BUSAM, D. A., CHAMPE, M., CENTER, A.,
CHAMPE, M., CHARTY, J. M., CHAVEZ, S. M., DE, P. L., FENG, W., FINE, B. C.,
FURN, P., GABOR, J. D., GILBERT, W. M., GLOCK, A., GOLF, F., GUTIEL, J. R.,
GUNZ, Q., HARRIS, M., HARTIS, N. L., HARVEY, P., HELMAN, J. J.,
HORNANDER, J. P., HOUK, J., HOSTIN, P., HOUSTON, K. A., HOWLAND, T. J.,
WEL, B. H., HEDMAN, J. D., JALALI, M., KALASH, I., KARPEN, S. H., KERR,
KENNEDY, J. A., KETCHEM, F. A., KIMMEL, R. E., KODIRA, S. E., KRAIT, C.,
KLEIN, S. G., KAP, E., LARSEN, L. K., LEE, Y., LELAND, A. A.,
LIT, J., LIZO, Z., LIANG, Y., LIU, X., LIU, X., MATTEI, R., MCINTOSH, J. C.,
MADON, M. J., MATHIAS, M., MATHIAS, S., MATHIAS, R. V., MATHIAS, C.,
MORRIS, J., MOSHREJI, A., MOUNT, S. M., MOY, M., MURPHY, B., MURPHY, L.,
MURPHY, L., NELSON, J. L., NELSON, E. F., NELSON, K. A., NELSON, K.,
NUSSEN, P. E., PACHECO, J. M., PACHECO, J. M., PACHECO, J. M., PACHECO, J. M.,
POLLARD, J. D., PURI, V., RABIN, M. G., RABIN, M. G., REMINGTON, K.,
SANDERS, P. D., SCHWARTZ, P., SKEN, H., SHU, P. C., SIDEN, KIMOS, I.,
STAMPSON, M., SKUPSKI, M. P., SMITH, T., SPIER, E., SPRADLING, A. C.,
VAL, C. M., ZHANG, Y., ZHANG, Y., ZHANG, Y., ZHANG, Y., ZHANG, Y.,
VENTER, E., WANG, A. H., WANG, X., WANG, Z. Y., WASSARMAN, D. A.,
WEISSBERG, J. M., WEISSBERG, J. M., WEISSBERG, J. M., WEISSBERG, J. M.,
WOLFEY, P. J., WU, Y., YANG, S., YANG, S., YANG, S., YANG, S.,
XIAO, J. N., ZHANG, W., ZHANG, X., ZHANG, X., ZHANG, X., ZHANG, X.,
GIBBS, R. A., MYERS, R. W., WATSON, G. M., and VENTER, J. C.,
The genome sequence of *Drosophila melanogaster*
Science, 287 (341), 216-219, (2000)
20196006
10731132
2 (bases 1 to 345026)
ANASTASIDES, P. G., BRADTON, R. C., ROGERS, Y. H., AND, H. J., BALDWIN, D.,
BRANDON, J., BESSON, K. Y., BUSAM, D. A., CHAMPE, M., CENTER, A.,
CHAMPE, M., CHARTY, J. M., CHAVEZ, S. M., DE, P. L., FENG, W., FINE, B. C.,
FURN, P., GABOR, J. D., GILBERT, W. M., GLOCK, A., GOLF, F., GUTIEL, J. R.,
GUNZ, Q., HARRIS, M., HARTIS, N. L., HARVEY, P., HELMAN, J. J.,
HORNANDER, J. P., HOUK, J., HOSTIN, P., HOUSTON, K. A., HOWLAND, T. J.,
WEL, B. H., HEDMAN, J. D., JALALI, M., KALASH, I., KARPEN, S. H., KERR,
KENNEDY, J. A., KETCHEM, F. A., KIMMEL, R. E., KODIRA, S. E., KRAIT, C.,
KLEIN, S. G., KAP, E., LARSEN, L. K., LEE, Y., LELAND, A. A.,
LIT, J., LIZO, Z., LIANG, Y., LIU, X., LIU, X., MATTEI, R., MCINTOSH, J. C.,
MADON, M. J., MATHIAS, M., MATHIAS, S., MATHIAS, R. V., MATHIAS, C.,
MORRIS, J., MOSHREJI, A., MOUNT, S. M., MOY, M., MURPHY, B., MURPHY, L.,
MURPHY, L., NELSON, J. L., NELSON, E. F., NELSON, K. A., NELSON, K.,
NUSSEN, P. E., PACHECO, J. M., PACHECO, J. M., PACHECO, J. M., PACHECO, J. M.,
POLLARD, J. D., PURI, V., RABIN, M. G., RABIN, M. G., REMINGTON, K.,
SANDERS, P. D., SCHWARTZ, P., SKEN, H., SHU, P. C., SIDEN, KIMOS, I.,
STAMPSON, M., SKUPSKI, M. P., SMITH, T., SPIER, E., SPRADLING, A. C.,
VAL, C. M., ZHANG, Y., ZHANG, Y., ZHANG, Y., ZHANG, Y., ZHANG, Y.,
VENTER, E., WANG, A. H., WANG, X., WANG, Z. Y., WASSARMAN, D. A.,
WEISSBERG, J. M., WEISSBERG, J. M., WEISSBERG, J. M., WEISSBERG, J. M.,
WOLFEY, P. J., WU, Y., YANG, S., YANG, S., YANG, S., YANG, S.,
XIAO, J. N., ZHANG, W., ZHANG, X., ZHANG, X., ZHANG, X., ZHANG, X.,
GIBBS, R. A., MYERS, R. W., WATSON, G. M., and VENTER, J. C.,
The genome sequence of *Drosophila melanogaster*
Science, 287 (341), 216-219, (2000)
20196006
10731132
2 (bases 1 to 345026)
ANASTASIDES, P. G., BRADTON, R. C., ROGERS, Y. H., AND, H. J., BALDWIN, D.,
BRANDON, J., BESSON, K. Y., BUSAM, D. A., CHAMPE, M., CENTER, A.,
CHAMPE, M., CHARTY, J. M., CHAVEZ, S. M., DE, P. L., FENG, W., FINE, B. C.,
FURN, P., GABOR, J. D., GILBERT, W. M., GLOCK, A., GOLF, F., GUTIEL, J. R.,
GUNZ, Q., HARRIS, M., HARTIS, N. L., HARVEY, P., HELMAN, J. J.,
HORNANDER, J. P., HOUK, J., HOSTIN, P., HOUSTON, K. A., HOWLAND, T. J.,
WEL, B. H., HEDMAN, J. D., JALALI, M., KALASH, I., KARPEN, S. H., KERR,
KENNEDY, J. A., KETCHEM, F. A., KIMMEL, R. E., KODIRA, S. E., KRAIT, C.,
KLEIN, S. G., KAP, E., LARSEN, L. K., LEE, Y., LELAND, A. A.,
LIT, J., LIZO, Z., LIANG, Y., LIU, X., LIU, X., MATTEI, R., MCINTOSH, J. C.,
MADON, M. J., MATHIAS, M., MATHIAS, S., MATHIAS, R. V., MATHIAS, C.,
MORRIS, J., MOSHREJI, A., MOUNT, S. M., MOY, M., MURPHY, B., MURPHY, L.,
MURPHY, L., NELSON, J. L., NELSON, E. F., NELSON, K. A., NELSON, K.,
NUSSEN, P. E., PACHECO, J. M., PACHECO, J. M., PACHECO, J. M., PACHECO, J. M.,
POLLARD, J. D., PURI, V., RABIN, M. G., RABIN, M. G., REMINGTON, K.,
SANDERS, P. D., SCHWARTZ, P., SKEN, H., SHU, P. C., SIDEN, KIMOS, I.,
STAMPSON, M., SKUPSKI, M. P., SMITH, T., SPIER, E., SPRADLING, A. C.,
VAL, C. M., ZHANG, Y., ZHANG, Y.,

JOURNAL Chem. 272 (35), 2497-2498 (1997)
 MEDLINE 97450964
 REFERENCE 2 (bases 1 to 983)
 AUTHORS Oka.T.
 JOURNAL Direct Submission

TITLE Subcloned (ORF 1997) to cDNA library, the
 Institute of Scientific and Industrial Research, 8-1 Mihodaka,
 Ibaraki, 305-8565, Japan (P. O. Box 1, 8-1 Mihodaka, Ibaraki,
 Tel: 81-6-879-8491, Fax: 81-6-875-5724)

FEATURES Location/Qualifiers

1..983
 /variant="Caenorhabditis elegans"
 /db_xref="taxon:6239"
 24..509
 /codon_start=1
 /product="VHA-2"
 /protein_id="RAA22596.1"
 /db_xref="GI:12467494"

CSMSVMEPELIMKSVIVIMAGLILNGLVAMVKEVTSACAYLNKGAHLAAG
 LTCGLGCGAGYATGIVGAGVRCIAAGRIEVMILILIFSEVILGYOMIVALLIGT
 S"

PolyA_site 983

BASE COUNT 234 a 245 c 192 g 216 t

ORIGIN

Alignment Scores:
 Prod. No.: 4,376-08 Length: 983
 Score: 18.00 Matches: 18
 Percent Similarity: 100.00% Conservat: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.78% Indels: 0
 DB: 3 Gaps: 0

US 09-727-770-2 (1-205) x AB000918 (1-983)

QY 115 GlyThrAlaIndInpAratLeuHeVaGlyMetIleLeuIleIleIleIle 132
 ID 441 GGAACAGTCAACAGCCACGGTCTTTCGCGAATGATCTTATCTTTC 452

RESULT 46
 AB009566 1421 bp 2556 bases 11503 INV 20 FEB 1999
 LOCUS Caenorhabditis elegans mRNA for Vha3 protein, complete cds.
 DEFINITION Caenorhabditis elegans mRNA for Vha3 protein, complete cds.
 VERSION AB009566
 KEYWORDS Vha3 protein

SOURCE Caenorhabditis elegans cDNA to mRNA, clone lib:Yuji Kohara
 uncloned cDNA library G363781291
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditidae; Rhabditidae; Rhabditidae; Rhabditidae.

REFERENCE 1 (sites)
 Oka.T., Yamamoto.R. and Futai.M.
 Multiple genes for vesicular-type ATPase proteolipids in
 Caenorhabditis elegans. A new gene, vha 3, has a distinct
 cell-specific distribution

JOURNAL J. Biol. Chem. 273 (35), 22570-22576 (1998)
 MEDLINE 98480481
 REFERENCE 2 (bases 1 to 1421)
 Oka.T.
 Direct Submission

TITLE Submitted (04 FEB 1997) to cDNA library, the
 Institute of Scientific and Industrial Research, 8-1 Mihodaka,
 Ibaraki, 305-8565, Japan (P. O. Box 1, 8-1 Mihodaka, Ibaraki,
 Tel: 81-6-879-8491, Fax: 81-6-875-5724)

FEATURES Location/Qualifiers

1..1421
 /variant="Caenorhabditis elegans"
 /db_xref="taxon:6239"
 24..509
 /codon_start=1
 /product="Vha3"

Feature_lib "Yuji Kohara unpublished cDNA library"

BASE COUNT 234 a 245 c 192 g 216 t

ORIGIN

Alignment Scores:
 Prod. No.: 5,696-08 Length: 1421
 Score: 18.00 Matches: 18
 Percent Similarity: 100.00% Conservat: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 8.78% Indels: 0
 DB: 3 Gaps: 0

US 09-727-770-2 (1-205) x AB009566 (1-1421)
 QY 115 GlyThrAlaIndInpAratLeuHeVaGlyMetIleLeuIleIleIleIle 132
 ID 441 GGAACAGTCAACAGCCACGGTCTTTCGCGAATGATCTTATCTTTC 454

RESULT 47

AB009566 3620 bp 2556 bases 11503 INV 26 APR 1999
 LOCUS Newb12
 DEFINITION Ascaris lumbricoidea trans spliced mRNA encoded protein gene 12, complete cds.

ACCESSION M33757.1 GI:159658
 KEYWORDS Ascaris lumbricoidea suam DNA, clone gene 12.
 SOURCE Ascaris lumbricoidea
 Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida;
 Ascarididae; Ascarididae; Ascarididae.

REFERENCE 1 (bases 1 to 3620)
 Exner,J.A. and Nilsson,T.W.
 Trans splicing of nematode pre-mRNA in vitro
 JOURNAL Unpublished (1990)
 REFERENCE 2 (bases 886 to 1149)
 Nilsson,T.W., Harney,P.A., Exner,J.A. and Nilsson,T.W.
 Trans splicing of nematode pre-messenger RNA in vitro
 CELL 62 (7), 1247-1255 (1990)
 MEDLINE 90304884
 PUBMED 2164760

COMMENT Draft entry and complete readable sequence for [1] kindly submitted
 by T.W.Nilsson, 18-APR-1990.
 [1] Author address: T.W.Nilsson,
 Case Western Reserve University,
 Department of Molecular Biology and Microbiology
 School of Medicine, 2119
 Abington Rd., Cleveland OH, 44106,
 Ph: (216) 368-3913, FAX: 216-468-4055.

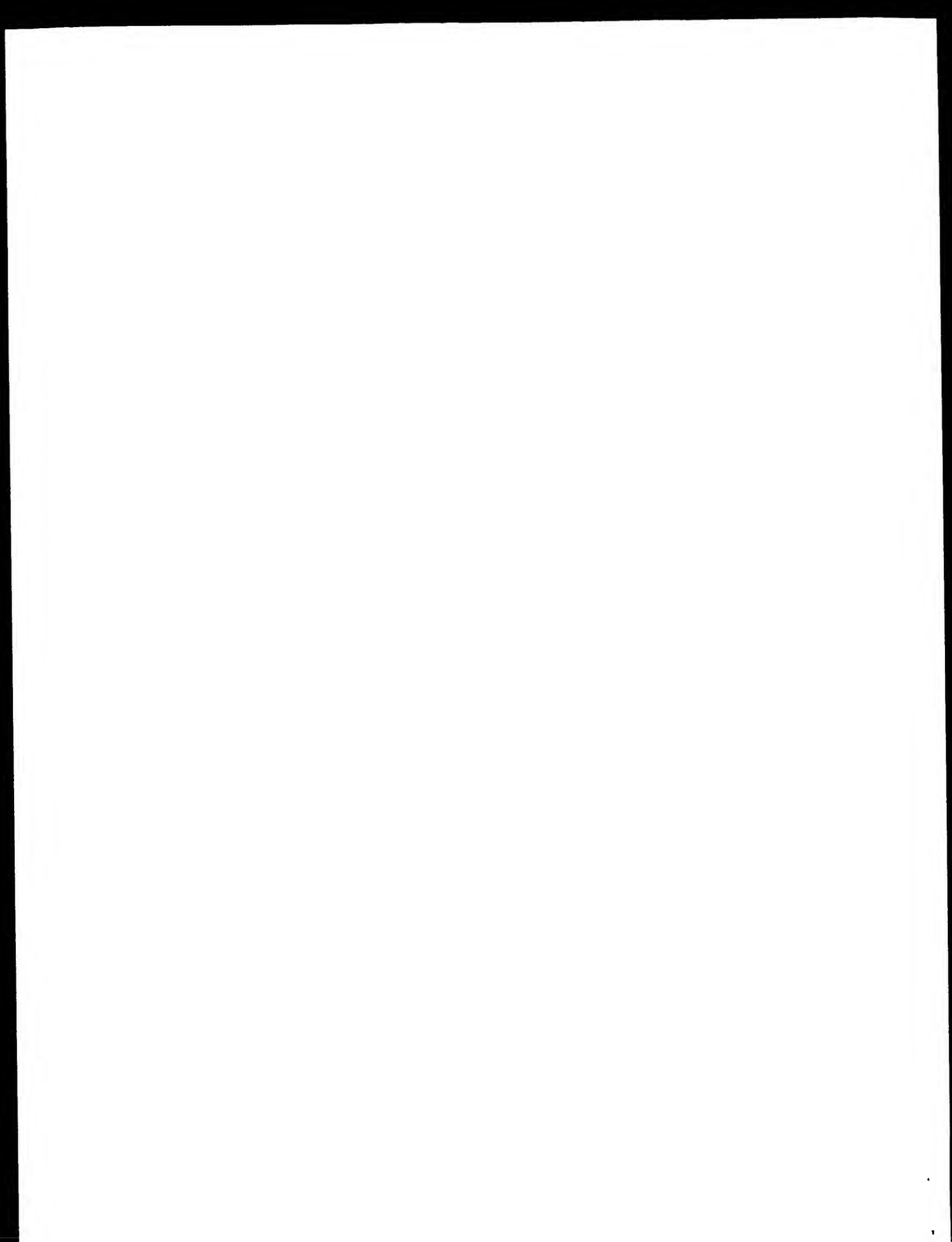
FEATURES

source

prim_transcript

prim_transcript

ORIGIN




```

PI /standard_name= "Single nucleotide polymorphism"
XX W-20194380-A2.
XX 13-DEC-2001.
PD
XX 31-MAY-2001; 2001WO-0517511.
XX
XX 02-JUN-2000; 2000US-208836P.
XX
XX 04-DEC-2000; 2000US-0727770.
XX
XX (APPL-) APPLERA CORP.
XX
XX Li Z, Wei M, Fetherly M, Fetherly M.
XX WPI: 2002 566438760.
XX P-PSDB: AA019455.
XX
XX New human transporter proteins related to the proton ATPase transporter
XX substantially for diagnosing and treating diseases mediated by the
XX transporter protein and for identifying modulators
XX
XX Claim 23: Fig 3; 70pp; English.
XX
XX The invention comprises the amino acid and coding sequence of a human
XX transporter protein, transporter proteins regulate many different
XX functions of a cell, including cell proliferation, differentiation and
XX signalling processes. Transporter proteins function by regulating the
XX flow of molecules such as ions and macromolecules into and out of cells.
XX The gene for the human transporter protein is located on chromosome 17.
XX The DNA and protein sequences of the invention are useful for identifying
XX agents that modulate or bind to the human transporter protein. The DNA
XX and protein sequences are useful for treating a disease or condition that
XX is mediated by a human transporter protein. The DNA and protein sequences
XX of the invention may also be used in the construction of transgenic
XX animals, pharmacogenomic analysis and for tissue typing. The present DNA
XX sequence encodes the human transporter protein of the invention.
XX
XX Sequence 6339 BP: 1613 A; 1626 C; 1609 G; 1491 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 1,90-196 Length: 6,339
XX Score: 205.00 Matches: 205
XX Percent Similarity: 100.00% Conservations: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-09-727-770 2 (1-205) x AA144994 (1-6339)
XX
XX 1 MetSerAsnAsnSerProGlyThrAlaLeuValPheThrIleGlyAlaLysAlaLys 20
XX
XX 3060 ATGTCCAAACAAACAGACCGGATATGCTTTGGTTCATCTGCGATGTAAGGCAAC 3059
XX
XX 21 MetValSerSerCysGlyAlaAlaCysGlyMetAlaLysAsnThrGlyIleMet 40
XX
XX 3060 ATGGTCTCAGTGGGTGGAGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3119
XX
XX 41 AlaMetSerValMetIleProGlyLeuIleHisMetLysSerIleIleProValValMet 60
XX
XX 3126 GCAATGCTGTGTCATGGGCAACAGTGTGATGCAATGCAATGCAATGCAATGCAAT 3179
XX
XX 42 AlaCysThrIleThrIleGlyThrValAlaValValPheValValValValVal 80
XX
XX 3180 GTGTATATCATCATCATCATATATGCTGATGCTGATGCTGATGCTGATGCTGAT 3239
XX
XX 43 AspAspAsnSerLeuLysSerSerPheLeuIleLeuValAlaLysCysGlyLeuAla 100
XX
XX 3240 GATGACACACAGTCTATAGCAATTCCTGCAATCTGCTGCTGCTGCTGCTGCTGCT 3299
XX
XX 44 AlaGlyPheAlaIleValIleValGlyAspThrGlyLysCysGlyThrAlaLeuIlePro 120
XX
XX 3300 ATGATGTTGGTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3359

```

```

QY 121 ArgLeuPheValGlyMetIleLeuIleLeuAlaPheValIleValIleLeuSerThr 140
XX
XX 3360 CGACTATTGTAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3419
XX
XX 141 LysClnProCysSerLysProThrSerHisArgIleArgCysLysAspHisProSerSer 160
XX
XX 3420 AAGCAGACCCCTCTCAAAACCCACCCAGTTCACAGCAATACGATGTAAAGACCCCTCTCTCA 3479
XX
XX 161 PheArgAlaValLysGlnProAspThrHisValLeuGlyLysSerThrPheSerValValAspLeu 180
XX
XX 3480 TTTCGGAAACAACACAGCCGACAGCCAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 3539
XX
XX 181 LysValValGlyCysPheCysValHisAlaIleAlaAlaValPheCysLeuPhePro 200
XX
XX 3540 CTAAGTGTATAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 3599
XX
XX 201 HisAlaValAspIle 205
XX
XX 4600 CATGCTGTGACATC 4614
XX
XX
XX RESULT 3
XX ARA46570
XX ID ARA46570 standard; DNA; 612 BP.
XX AC
XX ARA46570;
XX
XX C1 PUB 2002 (first entry)
XX
XX Human breast cell single exon nucleic acid probe #5265.
XX
XX Human, microarray, single exon probe, gene expression, breast;
XX disease; cancer; ss.
XX
XX Home sapiens.
XX
XX W200157271 A2.
XX
XX 09-AUG-2001.
XX
XX 10-JAN-2001; 2001WO-0800662.
XX
XX 04-FEB-2000; 2000US-0190312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 04-AUG-2000; 2000US-0632466.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236459.
XX
XX 04-OCT-2000; 2000GB-0021263.
XX
XX (MOLE ) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI, 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes
XX
XX Claim 4; SEQ ID NO 5265; 427pp + sequence listing; English.
XX
XX The invention relates to a spatially addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosis breast disease. Gene
XX expression analysis is useful for assessing the toxicity of chemical
XX agents on cells. The microarray of this invention presents a far greater

```


Db 303 GTCAGCTTTGGCATCATGATGAGAGAGAGAAAGAGGCAATGTCGCACTGGCCAGCAGCC 362
 QY 121 ArGleuPheValGlyMetIleLeuIleLeuIlePheAlaIleValLeuIleLeuSerThr 140
 Db 363 CGACTATTGTAGTAAATATCACTGATGCTTTTCCAAAGGAGTGCATCTCTTCACACA 422
 QY 141 LysGlnProLeuSerLysProThrSerHisValIleValArgCysLysAspHisProSerSer 160
 Db 423 AAGCAGCGCTCTCTCAAAAGCAGCAGTACAGAAATAGAGATGATAAGAGCAAGCTCTGCTCA 482
 QY 161 PheArgAsnLysGlnProAspThrHisValIleuGlySerIleProSerValValAspLeu 180
 Db 483 TTCGCAACAACACAGCTGACACAGCATGTCTCTGGCAGAGTGGACCTCATAGTTGATCTT 542
 QY 181 LeuSerValGlnCysProArgValHisArgLeuLeuValArgProCysProLeuProPro 200
 Db 543 CTAAGTGTACAGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 602
 QY 201 HisAlaVal 203
 Db 603 CATGCTGTG 611
 RESULT 5
 ABA31570
 ID ABA31570 standard; DNA: 612 BP.
 XX AC ABA31570;
 XX 23-JAN-2002 (first entry)
 DE Probe #10036 for gene expression analysis in human heart cell sample.
 XX Human gene expression; heart; microarray; vascular system; probe;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease, ss.
 XX Homo sapiens.
 XX WC200157274-A2.
 PN 09-AUG-2001.
 XX 10-JAN 2001: 2001WO-050666.
 XX 04 FEB-2000: 2000US-0180312.
 XX 26 MAY-2000: 2000US-0207456.
 XX 10-JUN-2000: 2000US-0608408.
 XX 04-AUG-2000: 2000US-0632366.
 XX 21-SEP-2000: 2000US-0234687.
 XX 27-SEP-2000: 2000US-0236359.
 XX 04-OCT-2000: 2000GB-0024263.
 (MOLECULAR DYNAMICS INC.)
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 WI: 2001-48899/53.
 Single exon nucleic acid probes for analyzing gene expression in human hearts.
 Claim 4: SEQ ID No 10036; 530pp; English.
 The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and/or prognosticating diseases of the human heart and vascular system, e.g., cardiovascular disease, hypertension, cardiac arrhythmia and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX

SQ Sequence 612 BP; 120 A; 195 C; 151 G; 146 T; 0 other;

Alignment Scores:

Pred. No.: 2,126-195 Length: 612
 Score: 203.00 Matches: 203
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.92% Indels: 0
 DB: 22 Gaps: 0

US-09-727-770-2 (1-205) x ABA31570 (1-612)

QY 1 MetSerAsnSerProGlnIleValAlaValPheThrIleSerGlyAlaMetAlaThr 20
 Db 3 ATGTCCAAACAAGCCGCGATATGTTTGTGTTTTCACATCTCGGCTGCTATGGCAAC 62
 QY 21 MetValSerSerLysLeuGlyAlaAlaGlyGlyMetAlaIleAsnGlyThrGlyIleMet 40
 Db 63 ATGGTCTCCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 122
 QY 41 AlaMetSerValMetIleProGlnLeuIleHisMetLysSerIleIleProValValMet 60
 Db 123 GCCATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
 QY 61 AlaGlyIleIleThrIleTyrGlyLeuValAlaValProAlaValProAlaValMet 80
 Db 183 GCTGTATCATCACCATCTATGGCTAGTGGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
 QY 91 AspAspAsnSerLeuTyrSerSerPheLeuGlnLeuGlyAlaValLeuSerGlyIleVal 100
 Db 243 CATACAAACAATATATAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 302
 QY 101 AlaGlyPheAlaIleValIleValGlyAspThrGlyLysCysGlyThrAlaGlnGlyPhe 120
 Db 303 GCTGGGCTTTGGCAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
 QY 121 ArgLeuPheValGlyMetIleLeuIleLeuIleAlaLysValLeuIleLeuSerThr 140
 Db 463 GCACTATTGTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 422
 QY 141 LysIleProLeuSerLysIleProThrSerHisValIleValArgCysLysAspHisProSer 160
 Db 423 AAGCAGCGCTCTCTCAAAAGCAGCAGTACAGAAATAGAGATGATAAGAGCAAGCTCTG 482
 QY 161 PheArgAsnLysGlnProAspThrHisValIleuGlySerIleProSerValValAspLeu 180
 Db 483 TTCGCAACAACACAGCTGACACAGCATGTCTCTGGCAGAGTGGACCTCATAGTTGATCT 542
 QY 181 LeuSerValGlnCysProArgValHisArgLeuLeuValArgProCysProLeuProPro 200
 Db 543 CTAAGTGTACAGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 602
 QY 201 HisAlaVal 203
 Db 603 CATGCTGTG 611

RESULT 6

AAK12891

ID AAK12891 standard; DNA: 612 BP.

XX AC AAK12891;

XX 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID No. 12891.

KW Human, brain expressed exon, gene expression analysis, probe;

KW microarray; Alzheimer's disease, multiple sclerosis, schizophrenia;
 KW epilepsy; cancer; ss.

01 30-JAN-2001; 2001W-US00665.
 XX 04-FEB-2000; 2000US-180312P.
 PR 27-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-060840P.
 PR 03-AUG-2000; 2000US-063236P.
 PR 21-SEP-2000; 2000US-244687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Peun SS, Hanzel DK, Chen W, Rank DK;
 XX
 DR WPI; 2002-114183/15.
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PI measure gene expression in human lung samples.
 PS
 PS Claim 1: SEQ ID No 56; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12-14 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 1987 open reading frames derived from the 12-14
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from nucleic acid
 CC of the eukaryote, and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from library of lung mRNA to nucleic acid probes
 CC having a fragment identical to the predicted exon; the probe is included
 CC in the above mentioned array, resulting in a detectable signal
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarray having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12-14 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF); the probes are used for gene
 CC expression analysis, and a hybridizing exon in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, neurofibromatosis,
 CC Niemann Pick disease, neurocysticercosis, sarcoidosis, pulmonary
 CC haemangiomas, pulmonary histiocytosis, lymphomatous teratoma,
 CC pulmonary alveolar proteinosis, Karman syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension,
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WPI at
 XX ftp.wpi.int/efc/patent/seqs/seqs-sequences
 XX
 XX Sequence: 448 BP; 79 A; 147 C; 112 G; 110 T; 0 other;
 SO

Alignment Scores:
 Prod. No.: 8-278-116 Length: 448
 Score: 124.00 Matches: 124
 Percent Similarity: 100.00% Conservation: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 60.45% Indels: 0
 Gaps: 0

US-09-727-770-2 (1-205) x AR000065 (1-448)

97 1 Met Ser Asn Asn Ser Pro Gly Tyr Ala Ala Val Phe His Ile Ser Gly Ala Met Ala Thr 20
 146 76 A T G C C A A A A C A G C C C C A G A T A G C T T T G G T T T C A C C A C T C G G T G C A T A G G C C A C C 145
 97 21 Met Val Ser Ser Gly Leu Gly Ala Ala Cys Gly Met Ala Cys Asn Gly Thr Gly Ile Met 40
 146 156 A G G T T C A A A A G G C C T G G G T A G G T A G G A 155
 97 41 Ala Met Ser Val Met Thr Pro Glu Leu Cys His Met Lys Ser Thr Cys Pro Val Val Met 60
 146 196 G C C A T A G C G C A G C A G C G C A G C A G C A G C A G C A G C A G C A G C A G C A G C A G C A G C A G C A G 255
 97 61 Ala Gly Ile Ile Thr Ile Thr Gly Gly Ala Ala Val Val Pro Pro Ala Asn Ser Leu Asn 80
 146 256 G C T G T A C A 315
 97 81 Asn Asp Asn Ser Leu Gly Ser Ser Thr Cys His Leu Leu Ala Gly Leu Ser Thr Gly Ala 100
 146 316 G A T G A C A 375
 97 101 Ala Gly Phe Ala Ile Val Ile Val Ile Val Gly Asp Thr Gly Lys Cys Gly Thr Ala Thr Ile Pro 120
 146 376 G A G G A T T T T G A A G T A T A G T G A G G A 435
 97 121 Arg Leu Phe Val 134
 146 436 G G A C T A T T T G C A 447
 RESULT 20
 ID ABL13071 standard; cDNA; 954 BP.
 XX ABL13071;
 XX 26 MAR 2002 (first entry)
 XX
 XX Bcr-philis melanoma derived P12/P16 probe (BL 14 to 6095).
 XX Bcr-philis developmental biology. Cell signalling, haematoides;
 XX pharmaceutical, gene, ss.
 XX Bcr-philis melanoma derived.
 XX
 XX W200171042 A2.
 XX 27 SEP 2001.
 XX 23-MAR-2001; 2001W-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 XX (PULKE) DE CORP NY.
 XX
 XX Venter SS, Adams M, Li FW, Myers EW;
 XX WPI; 2001-656865/75.
 XX P-US08; ABB68968.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1099 or more
 XX genes from Bcr-philis and for elucidating cell signalling and cell-cell
 XX interactions -
 XX
 XX Claim 1: SEQ ID NO 3695; 21pp - Sequence listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Bcr-philis. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL16511), expressed DNA
 CC sequences (ABL161840-ABL16175) and the encoded proteins

Sequence 8155 BP; 2313 A; 1636 C; 1785 G; 2421 T; 0 other;

Alignment Scores:
Pred. No.: 1,26-08 Length: 8155
Score: 19,00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 9.27% Indels: 0
Gaps: 24

US-09-727-770-2 (1-205) x AAB05726 (1 8155)

QY 115 GlyAlaGlnProArgGluPheValGlyMet IleLeuIleLeuIlePheAla 133
|||||
DB 165A GGCACAGACACAGCAGCAGCTGTCGGGCACTGATCTGATCTGCGGC 1598

RESUL1 34
AAV77655
ID AAV77655 standard; cDNA: 956 BP;
XX AAC77655;
XX AAC77655;
XX AAC77655;
DI 08-FEB-2001 (first entry)
XX Human cancer associated gene sequence SEQ ID NO:49.
DE Human; cancer associated gene; cancer antigen; detection; cancer;
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytotoxic; proliferative; vasculature; immunomodulator;
KW antidiabetic; antihypertensive; antihemorrhagic; antithrombotic; antiviral;
KW antidiabetic; antihypertensive; antihemorrhagic; antithrombotic; antiviral;
KW antidiabetic; antihypertensive; antihemorrhagic; antithrombotic; antiviral;
KW dermatological; neuroprotective; thrombolytic; coagulant; neoplastic;
KW dermatological; neuroprotective; thrombolytic; coagulant; neoplastic;
KW vasotrophic; antipsoriatic; antihypertensive; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW haematologic; thrombolytic; cardiovascular disorder; infection;
KW haematologic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss
XX Homo sapiens.
OS Homo sapiens.
PN W020005350-A1.
XX W020005350-A1.
PD 21-SEP-2000.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000W0-0505882.
XX 08-MAR-2000; 2000W0-0505882.
XX 12-MAR-1999; 9563 8124270.
XX 12-MAR-1999; 9563 8124270.
PA (HUMA) HUMAN CHR:06 3021 INC.
XX Rosen CA, Ruben SM;
PI Rosen CA, Ruben SM;
XX WPI: 2000-587533/55.
DR p-PSDB; AAB43446.
XX p-PSDB; AAB43446.
XX Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer
PT useful for treating or diagnosing e.g. cancer
PS Claim 1: Page 646-647; 2352pp; English.
XX Claim 1: Page 646-647; 2352pp; English.
XX AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43498 to AAB44249. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Examples of activities
CC include: cytostatic, proliferative; vasculature; immunomodulator;
CC antidiabetic; antihypertensive; antihemorrhagic; antithrombotic;
CC antidiabetic; antihypertensive; antihemorrhagic; antithrombotic;
CC antidiabetic; antihypertensive; antihemorrhagic; antithrombotic;
CC dermatological; neuroprotective; thrombolytic; coagulant; neoplastic;
CC dermatological; neuroprotective; thrombolytic; coagulant; neoplastic;
CC ameliorating medical conditions and diagnosing pathological conditions
CC ameliorating medical conditions and diagnosing pathological conditions
CC polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune

disorders, allergic reactions, graft versus host disease and organ
rejection, modulate haemostatic or thrombolytic activity, modulate
inflammation, cancers, cardiovascular disorders, neurological disease and
bacterial or viral infections. The peptides, nucleotides, antibodies,
agonists and antagonists may be also be used in drug screens. AAB78449 to
AAC78457 and AAB44240 represent sequences used in the exemplification of
the present invention.

XX Sequence 956 BP; 153 A; 440 C; 259 G; 207 T; 7 other;

Alignment Scores:
Pred. No.: 1,626-07 Length: 956
Score: 17,00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 8.29% Indels: 0
Gaps: 21

US-09-727-770-2 (1-205) x AAC77655 (1 956)

QY 117 AlaGlnProArgGluPheValGlyMet IleLeuIleLeuIlePheAla 133
|||||
DB 521 GGCACAGACAGCAGCAGCTGTCGGGCACTGATCTGATCTGCGGC 571

RESUL1 35
AAV90133
ID AAV90133 standard; cDNA: 109 BP;
XX AAC90133;
XX AAC90133;
DI 15-FEB-1999 (first entry)
XX EST clone DB190.
DE Human; secreted protein expressed in prokaryotic cells; haematopoietic;
KW Human; secreted protein expressed in prokaryotic cells; haematopoietic;
KW tissue growth; activity; inhibiting; chemokine; chemokinesis; haematologic;
KW receptor; liquid; thrombolytic; anti-inflammatory; adhesion; anti-tumour;
KW gene therapy; ss
XX Homo sapiens.
OS Homo sapiens.
PN W09845436-A2.
XX W09845436-A2.
PD 15-OCT-1998.
XX 15-OCT-1998.
XX 10-APR-1998; 98W0-0506955.
XX 10-APR-1998; 98W0-0506955.
XX 10-APR-1997; 97US-0838821.
XX 10-APR-1997; 97US-0838821.
PA (HUMA) GENETICS INC INC
XX Agostino MJ, Jacobs K, Lavallie EK, Mezey JM, Morberg D;
PI Racine LA, Spaulding V, Treacy M;
XX Racine LA, Spaulding V, Treacy M;
XX WPI: 1999-070077/06.
XX New polynucleotides encoding human secreted proteins derived from
CC e.g. human blood, kidney, foetal lung, placenta, testes, brain,
CC ovary, pituitary, retina and colon cDNA libraries.
PS Claim 1: Page 442; 618pp; English.
XX Claim 1: Page 442; 618pp; English.
XX The present sequence represents a human expressed sequence tag (EST).
CC The polynucleotide, which is a secreted EST, and the encoded protein
CC are predicted to have useful biological activities which would make
CC them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals, although no supporting data is
CC given. Suggested activities include: antifibrotic activity, immune
CC stimulating or suppressing activity, haematopoietic regulating
CC activity, tissue growth activity, antigenic activity,
CC chemokine/chemokinesis activity, haematologic and thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC adhesion/tumour invasion suppressor activity, tumour inhibition

CC activity. The polynucleotide may also be useful for gene therapy.

XX Sequence 136 BP, 24 A, 35 C, 26 G, 24 T, 0 other;

Alignment Scores:

Prod. No.: 0-000221 Length: 109
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.34% Indels: 0
Gaps: 0

US-09-727-770-2 (1-205) x AAS84403 (1-109)

QY 52 MetLysSerIleIleProValMetAlaClyThelle 64

LA 5 ATGAAGTCATCAACGAGTGGTCATGCGTGGATGATC 43

RESULT 36

AAS83403

IQ AAS84403 standard; cDNA: 541 BP.

XX AAS84403:

XX 14-FEB-2002 (first entry)

XX UNA encoded novel human diagnostic protein #19207.

XX Human: chromosome mapping; gene mapping; gene therapy; forensis;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens

XX W0200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-0208641.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0640167.

XX (HYPH-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YL

XX W01-2801-639462/73.

XX P-PSDB: ABL19215.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensis, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity.

XX Claim 1: SEQ ID No 19207; 104pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging or sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensis, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 541 BP, 83 A, 166 C, 134 G, 138 T, 20 other;

Alignment Scores:

Prod. No.: 0-00102 Length: 541
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.34% Indels: 0
Gaps: 0

US-09-727-770-2 (1-205) x AAS83403 (1-541)

QY 52 MetLysSerIleIleProValMetAlaClyThelle 64

DB 199 ATGAAGTCATCAACGAGTGGTCATGCGTGGATGATC 237

RESULT 37

ABL11647

ID ABL11647 standard; cDNA: 582 BP.

XX ABL11647:

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide, SEQ ID NO 29424.

XX Drosophila: developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX W0200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-0509231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWB, Myers EW;

XX W01-2801-639462/75.

XX P-PSDB: ABL67544.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions.

XX Claim 1: SEQ ID NO 29423; 21pp; Sequence listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins

XX (ABL57237-ABL72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 582 BP, 155 A, 153 C, 176 G, 150 T, 0 other;

Alignment Scores:

XX ABL11136;
 XX AC
 XX 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster; expressed polyprotein (40-590) to N^o 27890.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 XX
 XX W6200171042-A2.
 XX
 XX 27-SEP-2001
 XX
 XX 23-MAR-2001; 2001W0-0809241.
 XX
 XX 23-MAR-2000; 2000US-191647P.
 XX
 XX 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li FWB, Myers EW;
 XX
 XX WPI; 2001-656860/76
 XX P-FSUB; AB667033.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX PI genes from Drosophila and for elucidating cell signalling and cell-cell
 XX PI interactions.
 XX
 XX claim 1; SEQ ID N^o 27890; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX insecticides, therapeutics and pharmaceutical drugs. The invention
 XX discloses genomic DNA sequences (ABL16176-ABL16511), expressed DNA
 XX sequences (ABL161840-ABL16175) and the encoded proteins
 XX (AB67717-AB672072).
 XX
 XX The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at: http://wipo.int/patlib/seq/seq_14666.
 XX
 XX Sequence 5413 BP; 1369 A; 1484 C; 1406 G; 1156 T; 0 other;
 XX
 XX Alignment Scores:
 XX Prod. No.: 0.00922 Length: 5413
 XX Score: 13.00 Matches: 13
 XX Percent Similarity: 100.00% Conservations: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 6.44% Indels: 0
 XX DB: 23 Gaps: 0
 XX
 XX US 69-727-770-2 (1 205) x ABL11136 (1-5413)
 XX
 XX Q: 52 MetLysSerIleIleValValMetAlaIleIleIle 64
 XX
 XX DB: 352 ATGAAGTCCATTCATGAGTGGTATGAGTGGTGGTATCAT 418
 XX
 XX RESULT 41
 XX AA:50%05
 XX IL: AA:50%05 standard; DNA: 501 BP.
 XX
 XX AC: AAC50505;
 XX
 XX 18-SEP-2000 (first entry)
 XX
 XX Arabidopsis thaliana RNA treatment SEQ ID N^o 65090.
 XX
 XX Arabidopsis thaliana RNA treatment SEQ ID N^o 65090.
 XX
 XX Hybridization assay; genetic mapping; gene expression control.
 XX
 XX
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX Arabidopsis thaliana.
 XX
 XX UFL03425-A2.
 XX
 XX 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 990S-0121825.
 XX
 XX 05-MAR-1999; 990S-0123180.
 XX
 XX 09-MAR-1999; 990S-0123548.
 XX
 XX 23-MAR-1999; 990S-0125788.
 XX
 XX 25-MAR-1999; 990S-0126264.
 XX
 XX 29-MAR-1999; 990S-0126785.
 XX
 XX 01-APR-1999; 990S-0127462.
 XX
 XX 06-APR-1999; 990S-0128234.
 XX
 XX 08-APR-1999; 990S-0128714.
 XX
 XX 16-APR-1999; 990S-0128845.
 XX
 XX 19-APR-1999; 990S-0130077.
 XX
 XX 21-APR-1999; 990S-0130449.
 XX
 XX 23-APR-1999; 990S-0130510.
 XX
 XX 28-APR-1999; 990S-0130891.
 XX
 XX 30-APR-1999; 990S-0131449.
 XX
 XX 30-APR-1999; 990S-0132048.
 XX
 XX 30-APR-1999; 990S-0132407.
 XX
 XX 04-MAY-1999; 990S-0132484.
 XX
 XX 05-MAY-1999; 990S-0132485.
 XX
 XX 06-MAY-1999; 990S-0132486.
 XX
 XX 06-MAY-1999; 990S-0132487.
 XX
 XX 07-MAY-1999; 990S-0132863.
 XX
 XX 11-MAY-1999; 990S-0134256.
 XX
 XX 14-MAY-1999; 990S-0134218.
 XX
 XX 14-MAY-1999; 990S-0134219.
 XX
 XX 14-MAY-1999; 990S-0134221.
 XX
 XX 14-MAY-1999; 990S-0134370.
 XX
 XX 18-MAY-1999; 990S-0134768.
 XX
 XX 19-MAY-1999; 990S-0134941.
 XX
 XX 20-MAY-1999; 990S-0135124.
 XX
 XX 21-MAY-1999; 990S-0135353.
 XX
 XX 24-MAY-1999; 990S-0135629.
 XX
 XX 25-MAY-1999; 990S-0136021.
 XX
 XX 27-MAY-1999; 990S-0136342.
 XX
 XX 28-MAY-1999; 990S-0136782.
 XX
 XX 31-JUN-1999; 990S-0137233.
 XX
 XX 03-JUN-1999; 990S-0137528.
 XX
 XX 07-JUN-1999; 990S-0137502.
 XX
 XX 08-JUN-1999; 990S-0137724.
 XX
 XX 08-JUN-1999; 990S-0138094.
 XX
 XX 10-JUN-1999; 990S-0138540.
 XX
 XX 10-JUN-1999; 990S-0138847.
 XX
 XX 14-JUN-1999; 990S-0139119.
 XX
 XX 16-JUN-1999; 990S-0139452.
 XX
 XX 16-JUN-1999; 990S-0139453.
 XX
 XX 17-JUN-1999; 990S-0139492.
 XX
 XX 18-JUN-1999; 990S-0139454.
 XX
 XX 18-JUN-1999; 990S-0139455.
 XX
 XX 18-JUN-1999; 990S-0139456.
 XX
 XX 18-JUN-1999; 990S-0139457.
 XX
 XX 18-JUN-1999; 990S-0139458.
 XX
 XX 18-JUN-1999; 990S-0139459.
 XX
 XX 18-JUN-1999; 990S-0139460.
 XX
 XX 18-JUN-1999; 990S-0139461.
 XX
 XX 18-JUN-1999; 990S-0139462.
 XX
 XX 18-JUN-1999; 990S-0139463.
 XX
 XX 18-JUN-1999; 990S-0139750.
 XX
 XX 18-JUN-1999; 990S-0139763.
 XX
 XX 21-JUN-1999; 990S-0139817.
 XX
 XX 22-JUN-1999; 990S-0139899.
 XX
 XX 23-JUN-1999; 990S-0140353.
 XX
 XX 23-JUN-1999; 990S-0140354.

us-09-727-770-2.olil0p2n.rng

Tue Dec 31 14:03:13 2002

```

PR 24-JUN-1999; 990S-0140695;
PR 28-JUN-1999; 990S-0140824;
PR 29-JUN-1999; 990S-0140991;
PR 30-JUN-1999; 990S-0141287;
PR 01-JUL-1999; 990S-0141842;
PR 01-JUL-1999; 990S-0142154;
PR 02-JUL-1999; 990S-0142055;
PR 06-JUL-1999; 990S-0142490;
PR 08-JUL-1999; 990S-0142801;
PR 09-JUL-1999; 990S-0142920;
PR 12-JUL-1999; 990S-0142977;
PR 13-JUL-1999; 990S-0143542;
PR 14-JUL-1999; 990S-0143624;
PR 15-JUL-1999; 990S-0144005;
PR 16-JUL-1999; 990S-0144085;
PR 16-JUL-1999; 990S-0144086;
PR 19-JUL-1999; 990S-0144325;
PR 19-JUL-1999; 990S-0144331;
PR 19-JUL-1999; 990S-0144332;
PR 19-JUL-1999; 990S-0144333;
PR 19-JUL-1999; 990S-0144334;
PR 19-JUL-1999; 990S-0144335;
PR 20-JUL-1999; 990S-0144352;
PR 20-JUL-1999; 990S-0144632;
PR 20-JUL-1999; 990S-0144684;
PR 21-JUL-1999; 990S-0144814;
PR 21-JUL-1999; 990S-0145086;
PR 21-JUL-1999; 990S-0145088;
PR 22-JUL-1999; 990S-0145085;
PR 22-JUL-1999; 990S-0145087;
PR 22-JUL-1999; 990S-0145089;
PR 22-JUL-1999; 990S-0145192;
PR 22-JUL-1999; 990S-0145145;
PR 23-JUL-1999; 990S-0145218;
PR 23-JUL-1999; 990S-0145224;
PR 26-JUL-1999; 990S-0145276;
PR 27-JUL-1999; 990S-0145913;
PR 27-JUL-1999; 990S-0145918;
PR 27-JUL-1999; 990S-0145919;
PR 28-JUL-1999; 990S-0145951;
PR 02-AUG-1999; 990S-0146386;
PR 02-AUG-1999; 990S-0146388;
PR 02-AUG-1999; 990S-0146489;
PR 04-AUG-1999; 990S-0147038;
PR 04-AUG-1999; 990S-0147204;
PR 05-AUG-1999; 990S-0147302;
PR 05-AUG-1999; 990S-0147192;
PR 06-AUG-1999; 990S-0147260;
PR 06-AUG-1999; 990S-0147303;
PR 06-AUG-1999; 990S-0147416;
PR 09-AUG-1999; 990S-0147493;
PR 09-AUG-1999; 990S-0147935;
PR 10-AUG-1999; 990S-0148171;
PR 11-AUG-1999; 990S-0148341;
PR 12-AUG-1999; 990S-0148565;
PR 13-AUG-1999; 990S-0148684;
PR 16-AUG-1999; 990S-0149368;
PR 17-AUG-1999; 990S-0149175;
PR 18-AUG-1999; 990S-0149426;
PR 20-AUG-1999; 990S-0149722;
PR 20-AUG-1999; 990S-0149723;
PR 20-AUG-1999; 990S-0149929;
PR 23-AUG-1999; 990S-0149902;
PR 24-AUG-1999; 990S-0149930;
PR 25-AUG-1999; 990S-0150566;
PR 26-AUG-1999; 990S-0150884;
PR 27-AUG-1999; 990S-0151065;
PR 27-AUG-1999; 990S-0151066;
PR 27-AUG-1999; 990S-0151080;
PR 30-AUG-1999; 990S-0151303;
PR 31-AUG-1999; 990S-0151436;
PR 01-SEP-1999; 990S-0151930;

PR 07-SEP-1999; 990S-0152463;
PR 10-SEP-1999; 990S-0153070;
PR 13-SEP-1999; 990S-0153758;
PR 15-SEP-1999; 990S-0154018;
PR 16-SEP-1999; 990S-0154039;
PR 20-SEP-1999; 990S-0154779;
PR 22-SEP-1999; 990S-0155139;
PR 23-SEP-1999; 990S-0155486;
PR 24-SEP-1999; 990S-0155659;
PR 28-SEP-1999; 990S-0156458;
PR 29-SEP-1999; 990S-0156596;
PR 04-OCT-1999; 990S-0157117;
PR 05-OCT-1999; 990S-0157753;
PR 06-OCT-1999; 990S-0157865;
PR 07-OCT-1999; 990S-0158029;
PR 08-OCT-1999; 990S-0158232;
PR 12-OCT-1999; 990S-0158369;
PR 13-OCT-1999; 990S-0159293;
PR 13-OCT-1999; 990S-0159294;
PR 13-OCT-1999; 990S-0159295;
PR 14-OCT-1999; 990S-0159329;
PR 14-OCT-1999; 990S-0159330;
PR 14-OCT-1999; 990S-0159331;
PR 14-OCT-1999; 990S-0159637;
PR 14-OCT-1999; 990S-0159638;
PR 18-OCT-1999; 990S-0159584;
PR 21-OCT-1999; 990S-0160741;
PR 21-OCT-1999; 990S-0160767;
PR 21-OCT-1999; 990S-0160768;
PR 21-OCT-1999; 990S-0160770;
PR 21-OCT-1999; 990S-0160814;
PR 21-OCT-1999; 990S-0160815;
PR 22-OCT-1999; 990S-0160980;
PR 22-OCT-1999; 990S-0160981;
PR 22-OCT-1999; 990S-0160989;
PR 25-OCT-1999; 990S-0161404;
PR 25-OCT-1999; 990S-0161405;
PR 25-OCT-1999; 990S-0161406;
PR 26-OCT-1999; 990S-0161359;
PR 26-OCT-1999; 990S-0161460;
PR 26-OCT-1999; 990S-0161361;
PR 28-OCT-1999; 990S-0161920;
PR 28-OCT-1999; 990S-0161992;
PR 28-OCT-1999; 990S-0161993;
PR 29-OCT-1999; 990S-0162142;

Alignment Scores:
Pred. No.: 0.00969 501
Score: 12.00 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.85% Indels: 0
DB: 21 Gaps: 0

US 09 727-770-2 (1-205) x AAC50505 (1-501)
Qy 122 LeuphValGlyMetIleleuIleleuIleleuAla 133
Db 391 CTGTTTGGGGAATGATTCATCTTTGG 426

RESULT 42
AAC44322
Db AAC44322 standard; DNA: 575 BP.
XX
XX AAC44322;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SUB ID No: 42435;
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

```

XX US Arabidopsis thaliana.
XX EN EPI04405-A2.
XX PD 06-SEP-2000.
XX XX
XX 25-FEB-2000; 20000P-0301439.
XX
XX 25-FEB-1999; 990S-0121825.
XX 05-MAR-1999; 990S-0121180.
XX 19-MAR-1999; 990S-0124548.
XX 23-MAR-1999; 990S-0125788.
XX 25-MAR-1999; 990S-0126264.
XX 29-MAR-1999; 990S-0126785.
XX 01-APR-1999; 990S-0127462.
XX 06-APR-1999; 990S-0128234.
XX 08-APR-1999; 990S-0128714.
XX 16-APR-1999; 990S-0129845.
XX 19-APR-1999; 990S-0130077.
XX 21-APR-1999; 990S-0130449.
XX 23-APR-1999; 990S-0135199.
XX 24-APR-1999; 990S-0130891.
XX 28-APR-1999; 990S-0131449.
XX 30-APR-1999; 990S-0132048.
XX 31-APR-1999; 990S-0132407.
XX 04-MAY-1999; 990S-0132484.
XX 05-MAY-1999; 990S-0132485.
XX 06-MAY-1999; 990S-0132486.
XX 07-MAY-1999; 990S-0132487.
XX 11-MAY-1999; 990S-0132863.
XX 13-MAY-1999; 990S-0134256.
XX 14-MAY-1999; 990S-0134219.
XX 14-MAY-1999; 990S-0134221.
XX 14-MAY-1999; 990S-0134378.
XX 16-MAY-1999; 990S-0134768.
XX 19-MAY-1999; 990S-0134941.
XX 20-MAY-1999; 990S-0135124.
XX 21-MAY-1999; 990S-0135353.
XX 25-MAY-1999; 990S-0135629.
XX 27-MAY-1999; 990S-0136021.
XX 28-MAY-1999; 990S-0136792.
XX 01-JUN-1999; 990S-0136782.
XX 03-JUN-1999; 990S-0137222.
XX 04-JUN-1999; 990S-0137528.
XX 07-JUN-1999; 990S-0137502.
XX 08-JUN-1999; 990S-0137723.
XX 10-JUN-1999; 990S-0138094.
XX 13-JUN-1999; 990S-0138540.
XX 13-JUN-1999; 990S-0138847.
XX 14-JUN-1999; 990S-0139119.
XX 16-JUN-1999; 990S-0139452.
XX 16-JUN-1999; 990S-0139453.
XX 17-JUN-1999; 990S-0139492.
XX 18-JUN-1999; 990S-0139454.
XX 18-JUN-1999; 990S-0139455.
XX 18-JUN-1999; 990S-0139456.
XX 18-JUN-1999; 990S-0139457.
XX 18-JUN-1999; 990S-0139458.
XX 18-JUN-1999; 990S-0139459.
XX 18-JUN-1999; 990S-0139460.
XX 18-JUN-1999; 990S-0139461.
XX 18-JUN-1999; 990S-0139462.
XX 18-JUN-1999; 990S-0139463.
XX 18-JUN-1999; 990S-0139750.
XX 18-JUN-1999; 990S-0139764.
XX 21-JUN-1999; 990S-0139817.
XX 22-JUN-1999; 990S-0139899.
XX 23-JUN-1999; 990S-0140353.
XX 24-JUN-1999; 990S-0140354.
XX 24-JUN-1999; 990S-0140695.
XX 24-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144341.
PR 19-JUL-1999; 990S-0144342.
PR 19-JUL-1999; 990S-0144343.
PR 19-JUL-1999; 990S-0144344.
PR 19-JUL-1999; 990S-0144345.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144364.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 03-AUG-1999; 990S-0147234.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149436.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151940.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.

10 ABN98915 standard; DNA; 834 BP.
 XX
 AC ABN98915:
 XX
 XX
 DT 01-APR-2002 (first entry)
 XX
 DE Arabidopsis thaliana expressed polymorphic Id: SEQ ID NO 682.
 XX
 KW Arabidopsis thaliana; plant; insecticide; fumigant; transgenic; stress;
 KW disease; crop; thale cress; tolerance factor; insect; pathogen;
 XX nutrition; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 XX US200202281-A1.
 XX
 PD 21-FEB-2002.
 XX
 XX 25-JAN-2001; 2001US 0776415.
 XX
 PR 27-JAN-2000; 2000US-178472P.
 XX
 PA (GOLC/J) GRIPLATH T.
 PA (AMNY/J) AN Y.
 PA (HAMI/J) HAMILTON C M.
 PA (PRIC/J) PRICE J L.
 PA (RAIN/J) RAINES T M.
 PA (YUY/J) YU Y.
 PA (RAME/J) RAMEFAFA T S.
 PA (PAGE/J) PAGE A.
 PA (MATH/J) MATHW A V.
 PA (LEDF/J) LEDFORD B L.
 PA (WESS/J) WESSNER J P.
 PA (HAAS/J) HAAS W D.
 PA (GARC/J) GARCIA C A.
 PA (KRIC/J) KRICKER M.
 PA (SLAT/J) SLATER T.
 PA (DAVI/J) DAVIS K R.
 PA (ALIE/J) ALLEN K.
 PA (BOFF/J) BOFFMAN N.
 PA (HURB/J) HURBAN P.
 XX
 PI Gielach J, An Y, Hamilton CM, Price JB, Raines TM, Yu Y,
 PI Kaneoka JO, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Fickner M, Slater T, Davis JE, Allen K, Hoffman N,
 PI Hurban P;
 XX
 DE WPI: 2002-400741/42
 XX
 XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,
 PT producing compositions that modulate the expression or function of its
 PT encoded protein, and mapping functional regions of protein -
 XX
 PS claim 1: SEQ ID NO 683; 49pp - Sequence Listing; English.
 XX
 XX The invention relates to an Arabidopsis thaliana nucleic acid (I)
 CC comprising a sequence capable of hybridizing under stringent conditions
 CC to a sequence selected from any one of 599 sequences (ABN98233-ABN99231),
 CC given in the specification or its fragment. A polypeptide (II) encoded by
 CC (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
 CC genetically modified cell (IV) comprising an exogenous nucleic acid, is
 CC useful for screening a candidate agent for its biological effect. (I) is
 CC useful in identifying homologous or related genes, in producing
 CC compositions that modulate the expression or function of its encoded
 CC protein, mapping functional regions of the protein and in studying
 CC associated physiological pathways. (I) is also useful for the genetic
 CC manipulation of cells, particularly plant cells. (I) is also useful in
 CC screening assays of various plant strains to determine the strains that
 CC are best capable of withstanding a particular disease or environmental
 CC stress. (II) and (III) are useful for screening of biologically active
 CC agents, e.g. fungicides, insecticides, etc. for elucidating biochemical
 CC pathways. The screened agents are useful in improved methods of treating
 CC crops to prevent or treat disease. (II) are also useful in screening

CC programs to identify agents that mimic or enhance the action of tolerance
 CC factors. Such agents are useful in improved methods of treating crops to
 CC enhance their tolerance to environmental stress. (I) is also useful
 CC for enhancing or inhibiting production of a biosynthetic product in a
 CC plant. (III) is useful for identifying other mediators that may induce
 CC expression of proteins of interest, for establishing the extent to which
 CC any specific insect and/or pathogen is responsible for damage to a
 CC particular plant, for identifying other mediators that enhance or induce
 CC tolerance to environmental stress, for identifying factors involved in
 CC biosynthetic pathways of nutritional, commercial or medicinal value and
 CC for identifying productions of nutritional, commercial or medicinal
 CC value. (IV) is useful in the study of genetic function and regulation,
 CC for alteration of the cellular metabolism and for screening compounds
 CC that may affect the biological function of the gene or gene products.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 CC at SeqData 3317 10/25/99 00 00 00 00 00 00 00 00 00 00 00 00 00 00
 XX at SeqData 3317 10/25/99 00 00 00 00 00 00 00 00 00 00 00 00 00 00
 XX
 SO Sequence: 834 BP; 179 A; 164 C; 194 G; 297 T; 0 other;

 Alignment Scores:
 Pred. NO.: 0.0158 Length: 834
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.85% Indels: 0
 DB: 24 Gaps: 0

 US 09 727 770 2 (1-255) x ABN98915 (1-834)
 Q7 122 LeuPheValGlyMetIleuLeuIleuLeuPheAla 133
 |||||||||||||||||||||||||||||||||||||||
 D6 413 CMTITIGTGGAGCATTCCTTATCGTAAITTCGGA 448

 RESULT 45
 AAD03718
 ID AAD03718 standard; cDNA; 900 BP.
 AC
 XX AAD03718;
 XX
 DT 10-JUN-2001 (first entry)
 XX
 XX Tomato vacuolar protontranslocating ATPase (LVA-PI) cDNA.
 XX
 DE Tomato, vacuolar H⁺ ATPase, Lycopodium obscurum, Lycopodium obscurum, Lycopodium
 XX polygalacturonase, PG, LEXPG, cellulase, arabinosidase, LeABA;
 KW Lycopodium obscurum, Lycopodium obscurum, Lycopodium obscurum, Lycopodium
 KW LEXPG, LEXPG, LEXPG, LEXPG, LEXPG, LEXPG, LEXPG, LEXPG, LEXPG, LEXPG,
 KW seed germination; vacuolar protontranslocating ATPase, LVA-PI; ss.
 XX
 OS Lycopodium obscurum.
 XX
 XX Key Location/Qualifiers
 XX CDS 88..582
 XX /taq- a
 XX /product= "tomato vacuolar protontranslocating
 XX ATPase (LVA-PI)"
 XX
 XX W0200123530-A1.
 XX
 XX 05 APR 2001.
 XX
 XX 20-SEP-2000; 2000WO-US26884.
 XX
 XX 30-SEP-1999; 99US-0410191.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Researcher FJ, Chen F, Gatal D, LeuPhe B, Nishikawa H;
 XX WPI: 2001-266144/27.
 XX P-PSDB: AAE00419.
 XX


```

PR 21-JUL-1999; 99US-0145088
PR 22-JUL-1999; 99US-0145085
PR 22-JUL-1999; 99US-0145087
PR 22-JUL-1999; 99US-0145089
PR 22-JUL-1999; 99US-0145192
PR 23-JUL-1999; 99US-0145145
PR 23-JUL-1999; 99US-0145218
PR 23-JUL-1999; 99US-0145224
PR 26-JUL-1999; 99US-0145276
PR 27-JUL-1999; 99US-0145913
PR 27-JUL-1999; 99US-0145914
PR 27-JUL-1999; 99US-0145919
PR 28-JUL-1999; 99US-0145951
PR 02-AUG-1999; 99US-0146386
PR 02-AUG-1999; 99US-0146388
PR 02-AUG-1999; 99US-0146389
PR 03-AUG-1999; 99US-0147038
PR 04-AUG-1999; 99US-0147204
PR 04-AUG-1999; 99US-0147302
PR 05-AUG-1999; 99US-0147192
PR 05-AUG-1999; 99US-0147260
PR 06-AUG-1999; 99US-0147303
PR 06-AUG-1999; 99US-0147416
PR 09-AUG-1999; 99US-0147493
PR 09-AUG-1999; 99US-0147935
PR 10-AUG-1999; 99US-0148171
PR 11-AUG-1999; 99US-0148319
PR 12-AUG-1999; 99US-0148341
PR 13-AUG-1999; 99US-0148565
PR 13-AUG-1999; 99US-0148569
PR 16-AUG-1999; 99US-0149368
PR 17-AUG-1999; 99US-0149175
PR 18-AUG-1999; 99US-0149426
PR 20-AUG-1999; 99US-0149722
PR 20-AUG-1999; 99US-0149724
PR 20-AUG-1999; 99US-0149929
PR 20-AUG-1999; 99US-0149902
PR 24-AUG-1999; 99US-0149930
PR 25-AUG-1999; 99US-0150566
PR 26-AUG-1999; 99US-0150884
PR 27-AUG-1999; 99US-0151065
PR 27-AUG-1999; 99US-0151066
PR 27-AUG-1999; 99US-0151080
PR 30-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151438
PR 01-SEP-1999; 99US-0151930
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0153070
PR 13-SEP-1999; 99US-0153758
PR 15-SEP-1999; 99US-0154018
PR 16-SEP-1999; 99US-0154039
PR 20-SEP-1999; 99US-0154779
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155486
PR 24-SEP-1999; 99US-0155659
PR 28-SEP-1999; 99US-0156458
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157753
PR 06-OCT-1999; 99US-0157865
PR 07-OCT-1999; 99US-0158029
PR 08-OCT-1999; 99US-0158232
PR 12-OCT-1999; 99US-0158369
PR 13-OCT-1999; 99US-0159293
PR 13-OCT-1999; 99US-0159294
PR 14-OCT-1999; 99US-0159295
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638
PR 18-OCT-1999; 99US-0159584
PR 21-OCT-1999; 99US-0160741

```

```

PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160768
PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 22-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 22-OCT-1999; 99US-0160989
PR 25-OCT-1999; 99US-0161404
PR 25-OCT-1999; 99US-0161405
PR 25-OCT-1999; 99US-0161406
PR 26-OCT-1999; 99US-0161359
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161942
PR 28-OCT-1999; 99US-0161993
PR 29-OCT-1999; 99US-0162142

```

Aligned Scores:
 Pred. No.: 0.0171
 Score: 12.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 5.85%
 DB: 21

US-09-727-770-2 (1-205) x AAC50188 (1-907)
 Q7 122 :euPacValG;MetIleuAlaIlePheAla 133
 Db 493 CTTTCGTGGTAIGATCTTATCGATTCGTT 528

RESULT 47
 AAC37470
 ID AAC37430 standard; DNA: 911 BP.
 XX AAC37430;
 XX
 XX 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17350.
 XX
 KW Hybridisation assay, genetic mapping, gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence, ss.
 XX
 CS Arabidopsis thaliana.
 XX
 FN EP1033405-A2.
 XX
 PU 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 24-APR-1999; 99US-0140510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.

Tue Dec 31 14:03:13 2002

us-09-727-770-2.olil0p2n.rng

PR	05-MAY-1999;	990S-0132485.	PR	22-JUL-1999;	990S-0145087.
PR	06-MAY-1999;	990S-0132486.	PR	22-JUL-1999;	990S-0145088.
PR	06-MAY-1999;	990S-0132487.	PR	22-JUL-1999;	990S-0145192.
PR	07-MAY-1999;	990S-0132488.	PR	23-JUL-1999;	990S-0145145.
PR	11-MAY-1999;	990S-0134256.	PR	23-JUL-1999;	990S-0145218.
PR	14-MAY-1999;	990S-0134218.	PR	23-JUL-1999;	990S-0145224.
PR	14-MAY-1999;	990S-0134219.	PR	26-JUL-1999;	990S-0145276.
PR	14-MAY-1999;	990S-0134221.	PR	27-JUL-1999;	990S-0145913.
PR	14-MAY-1999;	990S-0134270.	PR	27-JUL-1999;	990S-0145918.
PR	18-MAY-1999;	990S-0134768.	PR	27-JUL-1999;	990S-0145919.
PR	19-MAY-1999;	990S-0134941.	PR	28-JUL-1999;	990S-0145951.
PR	20-MAY-1999;	990S-0135124.	PR	02-AUG-1999;	990S-0146386.
PR	21-MAY-1999;	990S-0135353.	PR	02-AUG-1999;	990S-0146388.
PR	24-MAY-1999;	990S-0135629.	PR	02-AUG-1999;	990S-0146489.
PR	25-MAY-1999;	990S-0136021.	PR	03-AUG-1999;	990S-0146489.
PR	27-MAY-1999;	990S-0136392.	PR	04-AUG-1999;	990S-0147038.
PR	28-MAY-1999;	990S-0136782.	PR	04-AUG-1999;	990S-0147204.
PR	01-JUN-1999;	990S-0137222.	PR	04-AUG-1999;	990S-0147302.
PR	03-JUN-1999;	990S-0137528.	PR	05-AUG-1999;	990S-0147192.
PR	04-JUN-1999;	990S-0137502.	PR	05-AUG-1999;	990S-0147260.
PR	07-JUN-1999;	990S-0137724.	PR	06-AUG-1999;	990S-0147404.
PR	08-JUN-1999;	990S-0138094.	PR	06-AUG-1999;	990S-0147416.
PR	10-JUN-1999;	990S-0138540.	PR	09-AUG-1999;	990S-0147494.
PR	10-JUN-1999;	990S-0138847.	PR	09-AUG-1999;	990S-0147935.
PR	14-JUN-1999;	990S-0139119.	PR	10-AUG-1999;	990S-0148171.
PR	16-JUN-1999;	990S-0139452.	PR	11-AUG-1999;	990S-0148319.
PR	16-JUN-1999;	990S-0139453.	PR	12-AUG-1999;	990S-0148341.
PR	17-JUN-1999;	990S-0139492.	PR	13-AUG-1999;	990S-0148565.
PR	18-JUN-1999;	990S-0139454.	PR	14-AUG-1999;	990S-0148684.
PR	18-JUN-1999;	990S-0139455.	PR	14-AUG-1999;	990S-0149368.
PR	18-JUN-1999;	990S-0139456.	PR	16-AUG-1999;	990S-0149175.
PR	18-JUN-1999;	990S-0139457.	PR	17-AUG-1999;	990S-0149175.
PR	18-JUN-1999;	990S-0139458.	PR	18-AUG-1999;	990S-0149426.
PR	18-JUN-1999;	990S-0139459.	PR	20-AUG-1999;	990S-0149722.
PR	18-JUN-1999;	990S-0139460.	PR	20-AUG-1999;	990S-0149723.
PR	18-JUN-1999;	990S-0139461.	PR	20-AUG-1999;	990S-0149929.
PR	18-JUN-1999;	990S-0139462.	PR	23-AUG-1999;	990S-0149902.
PR	18-JUN-1999;	990S-0139463.	PR	23-AUG-1999;	990S-0149930.
PR	18-JUN-1999;	990S-0139464.	PR	25-AUG-1999;	990S-0150566.
PR	18-JUN-1999;	990S-0139465.	PR	25-AUG-1999;	990S-0150884.
PR	18-JUN-1999;	990S-0139466.	PR	27-AUG-1999;	990S-0151065.
PR	18-JUN-1999;	990S-0139467.	PR	27-AUG-1999;	990S-0151066.
PR	18-JUN-1999;	990S-0139468.	PR	27-AUG-1999;	990S-0151080.
PR	18-JUN-1999;	990S-0139469.	PR	30-AUG-1999;	990S-0151303.
PR	18-JUN-1999;	990S-0139470.	PR	31-AUG-1999;	990S-0151438.
PR	18-JUN-1999;	990S-0139471.	PR	01-SEP-1999;	990S-0151930.
PR	18-JUN-1999;	990S-0139472.	PR	01-SEP-1999;	990S-0152363.
PR	18-JUN-1999;	990S-0139473.	PR	07-SEP-1999;	990S-0153070.
PR	18-JUN-1999;	990S-0139474.	PR	10-SEP-1999;	990S-0153758.
PR	18-JUN-1999;	990S-0139475.	PR	13-SEP-1999;	990S-0154018.
PR	18-JUN-1999;	990S-0139476.	PR	15-SEP-1999;	990S-0154039.
PR	18-JUN-1999;	990S-0139477.	PR	16-SEP-1999;	990S-0154779.
PR	18-JUN-1999;	990S-0139478.	PR	20-SEP-1999;	990S-0155139.
PR	18-JUN-1999;	990S-0139479.	PR	22-SEP-1999;	990S-0155486.
PR	18-JUN-1999;	990S-0139480.	PR	23-SEP-1999;	990S-0155659.
PR	18-JUN-1999;	990S-0139481.	PR	24-SEP-1999;	990S-0156458.
PR	18-JUN-1999;	990S-0139482.	PR	28-SEP-1999;	990S-0156596.
PR	18-JUN-1999;	990S-0139483.	PR	29-SEP-1999;	990S-0157117.
PR	18-JUN-1999;	990S-0139484.	PR	04-OCT-1999;	990S-0157753.
PR	18-JUN-1999;	990S-0139485.	PR	05-OCT-1999;	990S-0157865.
PR	18-JUN-1999;	990S-0139486.	PR	06-OCT-1999;	990S-0158029.
PR	18-JUN-1999;	990S-0139487.	PR	07-OCT-1999;	990S-0158232.
PR	18-JUN-1999;	990S-0139488.	PR	08-OCT-1999;	990S-0158369.
PR	18-JUN-1999;	990S-0139489.	PR	12-OCT-1999;	990S-0159293.
PR	18-JUN-1999;	990S-0139490.	PR	13-OCT-1999;	990S-0159294.
PR	18-JUN-1999;	990S-0139491.	PR	13-OCT-1999;	990S-0159295.
PR	18-JUN-1999;	990S-0139492.	PR	14-OCT-1999;	990S-0159429.
PR	18-JUN-1999;	990S-0139493.	PR	14-OCT-1999;	990S-0159430.
PR	18-JUN-1999;	990S-0139494.	PR	14-OCT-1999;	990S-0159631.
PR	18-JUN-1999;	990S-0139495.	PR	14-OCT-1999;	990S-0159637.
PR	18-JUN-1999;	990S-0139496.	PR	14-OCT-1999;	990S-0159638.
PR	18-JUN-1999;	990S-0139497.	PR	14-OCT-1999;	990S-0159584.
PR	18-JUN-1999;	990S-0139498.	PR	21-OCT-1999;	990S-0160741.
PR	18-JUN-1999;	990S-0139499.	PR	21-OCT-1999;	990S-0160767.
PR	18-JUN-1999;	990S-0139500.	PR	21-OCT-1999;	990S-0160768.

us-09-727-770-2.olil0p2n.rng

Tue Dec 31 14:03:13 2002

```

PR 22-JUL-1999; 990S-0145192
PR 23-JUL-1999; 990S-0145145
PR 23-JUL-1999; 990S-0145218
PR 23-JUL-1999; 990S-0145224
PR 26-JUL-1999; 990S-0145276
PR 27-JUL-1999; 990S-0145913
PR 27-JUL-1999; 990S-0145918
PR 27-JUL-1999; 990S-0145919
PR 28-JUL-1999; 990S-0145951
PR 02-AUG-1999; 990S-0146386
PR 02-AUG-1999; 990S-0146388
PR 02-AUG-1999; 990S-0146389
PR 03-AUG-1999; 990S-0147038
PR 04-AUG-1999; 990S-0147204
PR 04-AUG-1999; 990S-0147302
PR 05-AUG-1999; 990S-0147192
PR 05-AUG-1999; 990S-0147260
PR 06-AUG-1999; 990S-0147303
PR 06-AUG-1999; 990S-0147416
PR 09-AUG-1999; 990S-0147493
PR 09-AUG-1999; 990S-0147935
PR 10-AUG-1999; 990S-0148171
PR 11-AUG-1999; 990S-0148319
PR 12-AUG-1999; 990S-0148341
PR 13-AUG-1999; 990S-0148565
PR 13-AUG-1999; 990S-0148684
PR 16-AUG-1999; 990S-0149368
PR 17-AUG-1999; 990S-0149175
PR 17-AUG-1999; 990S-0149426
PR 18-AUG-1999; 990S-0149722
PR 20-AUG-1999; 990S-0149723
PR 20-AUG-1999; 990S-0149929
PR 23-AUG-1999; 990S-0149902
PR 23-AUG-1999; 990S-0149930
PR 23-AUG-1999; 990S-0150566
PR 25-AUG-1999; 990S-0150884
PR 25-AUG-1999; 990S-0151065
PR 27-AUG-1999; 990S-0151066
PR 27-AUG-1999; 990S-0151080
PR 27-AUG-1999; 990S-0151303
PR 30-AUG-1999; 990S-0151348
PR 31-AUG-1999; 990S-0151930
PR 01-SEP-1999; 990S-0152363
PR 07-SEP-1999; 990S-0153070
PR 10-SEP-1999; 990S-0153758
PR 13-SEP-1999; 990S-0154018
PR 15-SEP-1999; 990S-0154039
PR 16-SEP-1999; 990S-0154779
PR 20-SEP-1999; 990S-0155139
PR 22-SEP-1999; 990S-0155486
PR 23-SEP-1999; 990S-0155659
PR 24-SEP-1999; 990S-0156458
PR 28-SEP-1999; 990S-0156596
PR 29-SEP-1999; 990S-0157117
PR 04-OCT-1999; 990S-0157753
PR 05-OCT-1999; 990S-0157865
PR 06-OCT-1999; 990S-0158029
PR 07-OCT-1999; 990S-0158232
PR 08-OCT-1999; 990S-0158369
PR 12-OCT-1999; 990S-0159293
PR 13-OCT-1999; 990S-0159294
PR 13-OCT-1999; 990S-0159295
PR 13-OCT-1999; 990S-0159329
PR 14-OCT-1999; 990S-0159330
PR 14-OCT-1999; 990S-0159331
PR 14-OCT-1999; 990S-0159637
PR 14-OCT-1999; 990S-0159638
PR 18-OCT-1999; 990S-0159584
PR 21-OCT-1999; 990S-0160741
PR 21-OCT-1999; 990S-0160767
PR 21-OCT-1999; 990S-0160768
PR 21-OCT-1999; 990S-0160770
PR 21-OCT-1999; 990S-0160814

PR 21-OCT-1999; 990S-0160815
PR 22-OCT-1999; 990S-0160980
PR 22-OCT-1999; 990S-0160981
PR 22-OCT-1999; 990S-0160989
PR 25-OCT-1999; 990S-0161404
PR 25-OCT-1999; 990S-0161405
PR 25-OCT-1999; 990S-0161406
PR 26-OCT-1999; 990S-0161359
PR 26-OCT-1999; 990S-0161360
PR 26-OCT-1999; 990S-0161361
PR 26-OCT-1999; 990S-0161920
PR 28-OCT-1999; 990S-0161992
PR 28-OCT-1999; 990S-0161993
PR 28-OCT-1999; 990S-0162142

Alignment Scores:
Pred. No.: 0.0172
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.85%
DR: 21

US-09-727-770-2 (1-205) x AAC38627 (1-911)
QY 122 LeupheValglyMetIleLeuIleLeuIleLeuIlePheAla 133
DB 515 CTCTTTGTTGGGATGATCTTATCCCTTATTTCGCA 551

RESULT 49
AAC47752
ID AAC47752 standard; DNA: 977 BP.
AC AAC47752:
XX
XX 18-OCT-2000 (first entry)
DE Zea mays DNA fragment SEQ ID No: 54986.
XX
XX Hybridization assay; acnet ic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic;
XX pathway; promoter; termination sequence; curn; ss.
XX Zea mays subsp. mays.
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301449.
XX
XX 25-FEB-1999; 990S-0121825.
XX 05-MAR-1999; 990S-0123180.
XX 09-MAR-1999; 990S-0123548.
XX 23-MAR-1999; 990S-0125788.
XX 25-MAR-1999; 990S-0126264.
XX 29-MAR-1999; 990S-0126785.
XX 01-APR-1999; 990S-0127462.
XX 06-APR-1999; 990S-0128234.
XX 08-APR-1999; 990S-0128714.
XX 16-APR-1999; 990S-0129845.
XX 19-APR-1999; 990S-0130077.
XX 21-APR-1999; 990S-0130449.
XX 23-APR-1999; 990S-0130510.
XX 23-APR-1999; 990S-0130891.
XX 28-APR-1999; 990S-0131449.
XX 30-APR-1999; 990S-0132048.
XX 04-MAY-1999; 990S-0132407.
XX 05-MAY-1999; 990S-0132484.
XX 06-MAY-1999; 990S-0132486.
XX 06-MAY-1999; 990S-0132487.
XX 07-MAY-1999; 990S-0132863.

```

PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0144218.
PR 14-MAY-1999; 990S-0144219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139139.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139464.
PR 18-JUN-1999; 990S-0139465.
PR 18-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0143972.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144362.
PR 20-JUL-1999; 990S-0144384.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 06-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148441.
PR 13-AUG-1999; 990S-0148565.
PR 14-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 23-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 26-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155559.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 14-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159340.
PR 14-OCT-1999; 990S-0159341.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159648.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160940.


```
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145113.
PR 27-JUL-1999; 99US-0145118.
PR 27-JUL-1999; 99US-0145119.
PR 28-JUL-1999; 99US-0145151.
PR 02-AUG-1999; 99US-0146486.
PR 02-AUG-1999; 99US-0146488.
PR 02-AUG-1999; 99US-0146489.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147402.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147403.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147494.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148441.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149468.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151403.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152463.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 26-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157353.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158469.
PR 13-OCT-1999; 99US-0159233.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159429.
PR 14-OCT-1999; 99US-0159430.
PR 14-OCT-1999; 99US-0159431.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
```

```
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Alignment Scores:

Pred. No.:	0.0197	Length:	1052
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.85%	Indels:	0
Dr:	21	Gaps:	0

US 09 727 770-2 (1 205) x AN347753 (1-1052)

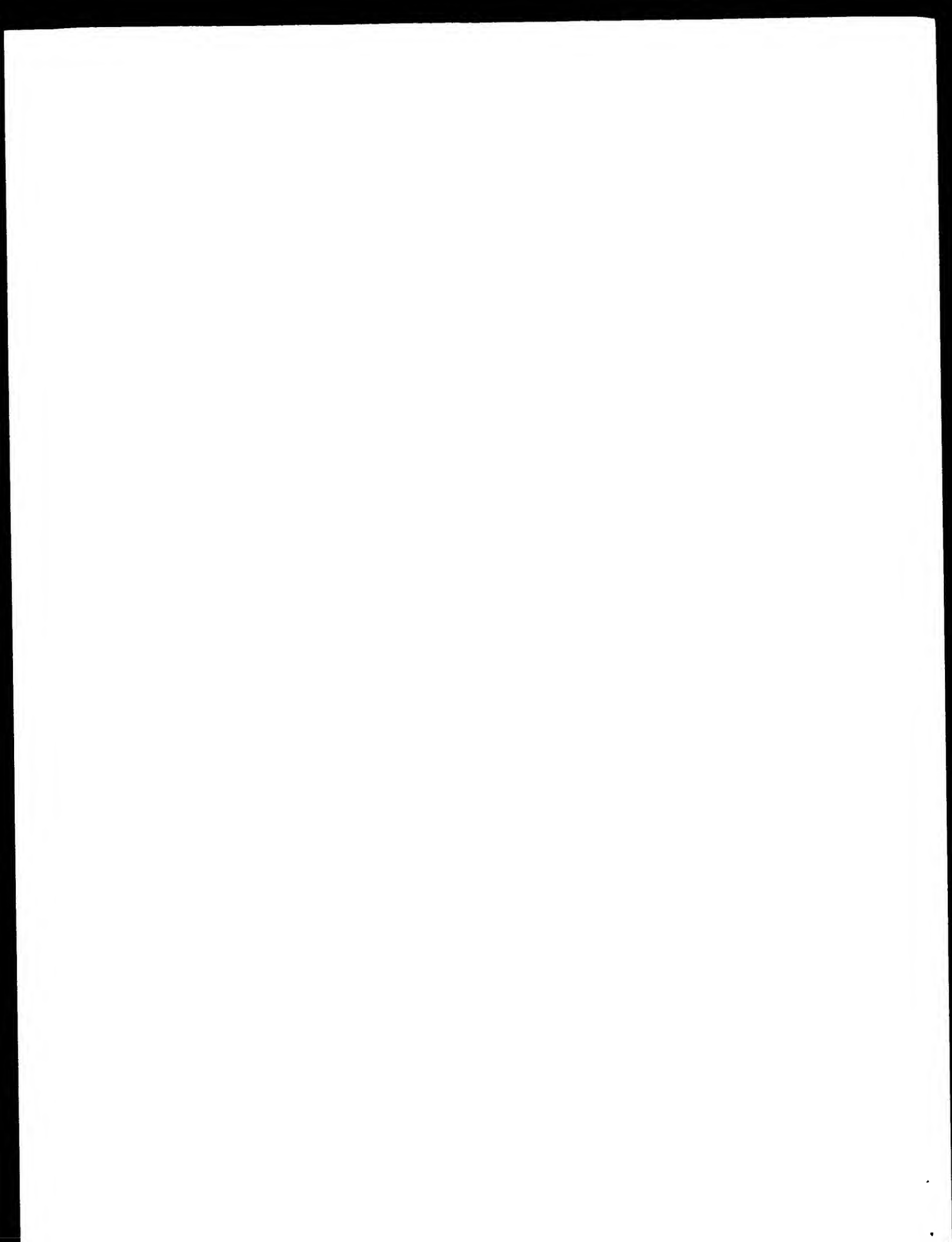
C7 122 AcPheValGlyMetIleuIleuIleuIleuPheAla 133

|||||

DB 569 TTTATTGTCATGATGTCATGTCATTTTCT 604

Search completed: December 28, 2002, 12:26:41

Job time : 268 secs



Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the full sequence distribution.

Genome version 5.1.3
Copyright (c) 1999-2002 Compugen Ltd

EM protein - nucleic search, using frame_plus_p2b model

Run on: December 28, 2002, 12:19:46 : Search time 1950 seconds
(without alignments)
17821422 Million cell updates/sec

Title: US-09-727-770-2
Perfect score: 205
Sequence: 1 MSKSTSTVATVITAGAGAA1.....AAVHMLAKGCLLTHAVVIL 265

Scoring table: OLIGO
Xgapop 60.0 : Xgapex 60.0
Ygapop 60.0 : Ygapex 60.0
Fgapop 6.0 : Fgapex 7.0
Delop 6.0 : Delext 7.0

Scored at: 16154065 cells, 8297743376 residues

Word size: 10

Total number of hits satisfying chosen parameters: 2489

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Command line parameters:

-MODEL frame_plus_p2b.model -DEV x11
-O 1 -N 1000000000 -S 1000000000 -T 1000000000 -C 1000000000
-DB EST -START -STOP -SUFFIX -PREFIX -MINMATCH 100 -MAXLEN 1000
-LOCAL 100 -END -1 -MATRIX -C 1000 -B 1000 -C 1000 -D 1000
-N 1000000000 -S 1000000000 -T 1000000000 -C 1000000000
-USER HPC973770 -N 1000000000 -S 1000000000 -T 1000000000
-W 1000000000 -S 1000000000 -T 1000000000 -C 1000000000
-WARN TIME 01-30 -THRESHOLD 1 -XGAPOP-60 -XGAPEX-60 -DELOP-60 -DELXT-60
-YGAPOP-60 -YGAPEX-60 -DELOP-60 -DELXT-60

Database :

- 1: em_estdb.*
- 2: em_estdb.*
- 3: em_estdb.*
- 4: em_estdb.*
- 5: em_estdb.*
- 6: em_estdb.*
- 7: em_estdb.*
- 8: em_estdb.*
- 9: db_estdb.*
- 10: db_estdb.*
- 11: db_estdb.*
- 12: db_estdb.*
- 13: db_estdb.*
- 14: db_estdb.*
- 15: em_estdb.*
- 16: em_estdb.*
- 17: db_estdb.*
- 18: em_estdb.*
- 19: em_estdb.*
- 20: em_estdb.*
- 21: em_estdb.*
- 22: em_estdb.*
- 23: em_estdb.*
- 24: em_estdb.*
- 25: em_estdb.*
- 26: em_estdb.*
- 27: em_estdb.*

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	20	9.8	629	9	AA579923	AA579923 AK09003.8
2	19	9.3	149	12	BF819273	BF819273 RC4-C1019
3	19	9.3	189	12	BF819278	BF819278 RC4-C1019
4	19	9.3	203	9	AL585139	AL585139 AL585139
5	14	9.3	228	14	BC333838	BC333838 BC333838
6	19	9.3	247	9	AA390838	AA390838 BC667404.F
7	19	9.3	306	13	BI768119	BI768119 BC3056543
8	19	9.3	326	13	BM280203	BM280203 BC3056543
9	19	9.3	339	14	BC305654	BC305654 BC305654
10	19	9.3	346	13	BI768119	BI768119 BC305654
11	19	9.3	346	14	BC305654	BC305654 BC305654
12	19	9.3	350	14	BM887954	BM887954 TMT1699.H
13	19	9.3	357	9	AI306652	AI306652 RC-BL122-
14	19	9.3	368	14	BQ371457	BQ371457 RC-BL122-
15	19	9.3	382	14	BQ333840	BQ333840 RC6-MT006
16	19	9.3	398	14	BQ372831	BQ372831 RC1-F1013
17	19	9.3	404	14	BQ333846	BQ333846 RC6-MT006
18	19	9.3	404	14	BQ371460	BQ371460 RC6-MT006
19	19	9.3	410	14	BQ372826	BQ372826 RC1-F1013
20	19	9.3	411	13	BM444137	BM444137 11L4A5.ab
21	19	9.3	412	12	BF601811	BF601811 266822.MA
22	19	9.3	412	14	BQ333842	BQ333842 RC2-BN033
23	19	9.3	416	14	BQ333842	BQ333842 RC6-MT006
24	19	9.3	417	14	BM836778	BM836778 K-PST0098
25	19	9.3	419	14	BQ372823	BQ372823 RC1-F1013
26	19	9.3	429	12	BF114171	BF114171 AK116117
27	19	9.3	432	12	BF601811	BF601811 266822.MA
28	19	9.3	431	10	AW441197	AW441197 RC200005A
29	19	9.3	431	10	AW441197	AW441197 RC200005A
30	19	9.3	439	10	BF777782	BF777782 RC777782
31	19	9.3	446	14	BQ333846	BQ333846 RC6-MT006
32	19	9.3	450	14	BQ371460	BQ371460 RC6-MT006
33	19	9.3	453	9	AI125755	AI125755 RC1-F1013
34	19	9.3	458	12	BF295283	BF295283 RC295283
35	19	9.3	462	10	BE430868	BE430868 166118.HA
36	19	9.3	464	13	BM619687	BM619687 170006874
37	19	9.3	465	10	BE654182	BE654182 UI-M-AN1
38	19	9.3	465	13	BF757330	BF757330 604029546
39	19	9.3	474	12	BF601869	BF601869 2669633.MA
40	19	9.3	487	14	BQ363476	BQ363476 QV9-ST021
41	19	9.3	489	12	BF706928	BF706928 281593.MA
42	19	9.3	491	10	BE276276	BE276276 601144.81
43	19	9.3	496	10	AW768111	AW768111 da71b02.Y
44	19	9.3	501	12	BE897736	BE897736 601439274
45	19	9.3	502	10	BE480867	BE480867 166117.BA
46	19	9.3	502	12	BF152692	BF152692 602525905
47	19	9.3	503	9	AI477910	AI477910 4657409.Y
48	19	9.3	504	12	BF245129	BF245129 RC54C09.Y
49	19	9.3	507	13	BI910735	BI910735 603070423
50	19	9.3	512	12	BC345676	BC345676 603070423
51	19	9.3	513	9	AI644755	AI644755 AL644755
52	19	9.3	513	13	BI768119	BI768119 BC3056543
53	19	9.3	514	9	AI534295	AI534295 SD07076.5
54	19	9.3	514	13	BI216891	BI216891 RE24173.5
55	19	9.3	515	9	AI541925	AI541925 SD07890.5
56	19	9.3	518	13	BM614099	BM614099 170006871
57	19	9.3	522	9	AI662267	AI662267 GH01415.5
58	19	9.3	522	9	AI518898	AI518898 1035744.5
59	19	9.3	522	9	AA264569	AA264569 1035744.5
60	19	9.3	523	9	AI264012	AI264012 LP05074.5
61	19	9.3	523	10	BE478394	BE478394 162630.BA
62	19	9.3	523	12	BC553393	BC553393 da325008
63	19	9.3	526	9	AI518612	AI518612 1039224.5
64	19	9.3	526	12	BC515254	BC515254 da366969

357	19	9.3	430	19	9.3	674	13	B1914587	B1914587
358	19	9.3	431	19	9.3	675	13	B629943	B629943
359	19	9.3	432	19	9.3	676	13	AV553614	AV553614
360	19	9.3	433	19	9.3	677	13	AV599044	AV599044
361	19	9.3	434	19	9.3	678	13	B629943	B629943
362	19	9.3	435	19	9.3	679	13	B629943	B629943
363	19	9.3	436	19	9.3	680	13	B629943	B629943
364	19	9.3	437	19	9.3	681	13	B629943	B629943
365	19	9.3	438	19	9.3	682	13	B629943	B629943
366	19	9.3	439	19	9.3	683	13	B629943	B629943
367	19	9.3	440	19	9.3	684	13	B629943	B629943
368	19	9.3	441	19	9.3	685	13	B629943	B629943
369	19	9.3	442	19	9.3	686	13	B629943	B629943
370	19	9.3	443	19	9.3	687	13	B629943	B629943
371	19	9.3	444	19	9.3	688	13	B629943	B629943
372	19	9.3	445	19	9.3	689	13	B629943	B629943
373	19	9.3	446	19	9.3	690	13	B629943	B629943
374	19	9.3	447	19	9.3	691	13	B629943	B629943
375	19	9.3	448	19	9.3	692	13	B629943	B629943
376	19	9.3	449	19	9.3	693	13	B629943	B629943
377	19	9.3	450	19	9.3	694	13	B629943	B629943
378	19	9.3	451	19	9.3	695	13	B629943	B629943
379	19	9.3	452	19	9.3	696	13	B629943	B629943
380	19	9.3	453	19	9.3	697	13	B629943	B629943
381	19	9.3	454	19	9.3	698	13	B629943	B629943
382	19	9.3	455	19	9.3	699	13	B629943	B629943
383	19	9.3	456	19	9.3	700	13	B629943	B629943
384	19	9.3	457	19	9.3	701	13	B629943	B629943
385	19	9.3	458	19	9.3	702	13	B629943	B629943
386	19	9.3	459	19	9.3	703	13	B629943	B629943
387	19	9.3	460	19	9.3	704	13	B629943	B629943
388	19	9.3	461	19	9.3	705	13	B629943	B629943
389	19	9.3	462	19	9.3	706	13	B629943	B629943
390	19	9.3	463	19	9.3	707	13	B629943	B629943
391	19	9.3	464	19	9.3	708	13	B629943	B629943
392	19	9.3	465	19	9.3	709	13	B629943	B629943
393	19	9.3	466	19	9.3	710	13	B629943	B629943
394	19	9.3	467	19	9.3	711	13	B629943	B629943
395	19	9.3	468	19	9.3	712	13	B629943	B629943
396	19	9.3	469	19	9.3	713	13	B629943	B629943
397	19	9.3	470	19	9.3	714	13	B629943	B629943
398	19	9.3	471	19	9.3	715	13	B629943	B629943
399	19	9.3	472	19	9.3	716	13	B629943	B629943
400	19	9.3	473	19	9.3	717	13	B629943	B629943
401	19	9.3	474	19	9.3	718	13	B629943	B629943
402	19	9.3	475	19	9.3	719	13	B629943	B629943
403	19	9.3	476	19	9.3	720	13	B629943	B629943
404	19	9.3	477	19	9.3	721	13	B629943	B629943
405	19	9.3	478	19	9.3	722	13	B629943	B629943
406	19	9.3	479	19	9.3	723	13	B629943	B629943
407	19	9.3	480	19	9.3	724	13	B629943	B629943
408	19	9.3	481	19	9.3	725	13	B629943	B629943
409	19	9.3	482	19	9.3	726	13	B629943	B629943
410	19	9.3	483	19	9.3	727	13	B629943	B629943
411	19	9.3	484	19	9.3	728	13	B629943	B629943
412	19	9.3	485	19	9.3	729	13	B629943	B629943
413	19	9.3	486	19	9.3	730	13	B629943	B629943
414	19	9.3	487	19	9.3	731	13	B629943	B629943
415	19	9.3	488	19	9.3	732	13	B629943	B629943
416	19	9.3	489	19	9.3	733	13	B629943	B629943
417	19	9.3	490	19	9.3	734	13	B629943	B629943
418	19	9.3	491	19	9.3	735	13	B629943	B629943
419	19	9.3	492	19	9.3	736	13	B629943	B629943
420	19	9.3	493	19	9.3	737	13	B629943	B629943
421	19	9.3	494	19	9.3	738	13	B629943	B629943
422	19	9.3	495	19	9.3	739	13	B629943	B629943
423	19	9.3	496	19	9.3	740	13	B629943	B629943
424	19	9.3	497	19	9.3	741	13	B629943	B629943
425	19	9.3	498	19	9.3	742	13	B629943	B629943
426	19	9.3	499	19	9.3	743	13	B629943	B629943
427	19	9.3	500	19	9.3	744	13	B629943	B629943
428	19	9.3	501	19	9.3	745	13	B629943	B629943
429	19	9.3	502	19	9.3	746	13	B629943	B629943

504	19	9.3	726	13	R1146793	6020911873	576	19	9.3	833	13	R1555191	6012461336
504	19	9.3	726	14	R0347047	AI-2F11-E	577	19	9.3	835	12	R1526968	6020702437
505	19	9.3	727	14	R0301440	01-E-D21-E	578	19	9.3	837	9	A0050682	A0050682
506	19	9.3	729	12	R0942635	16C7-GNA	579	19	9.3	837	14	R1246866	602095959
507	19	9.3	730	10	R0294765	6019408020	580	19	9.3	841	14	R0743163	602031163
508	19	9.3	730	14	R1409789	602062949	581	19	9.3	846	12	R1204276	602030101
509	19	9.3	730	14	R0676657	170036871	582	19	9.3	846	11	R0318721	A00318721
510	19	9.3	731	14	R1487079	R07058415	583	19	9.3	848	9	A0167464	A0167464
511	19	9.3	731	14	R0221173	02-E-E-E01	584	19	9.3	852	14	R0687546	A00687546
512	19	9.3	732	9	AA567668	HE15015.5	585	19	9.3	854	12	R1244306	602042218
513	19	9.3	733	14	R1154161	602002654	586	19	9.3	854	12	R1907237	601804854
514	19	9.3	735	14	R0377932	01-M-EPO-	587	19	9.3	855	14	R1866910	602087067
515	19	9.3	737	9	AI557998	A1557998	588	19	9.3	856	12	R1338911	602080634
516	19	9.3	743	14	R0579218		589	19	9.3	859	12	R122746	6020761274
517	19	9.3	742	10	R0563326		590	19	9.3	862	13	R1453128	6020713125
518	19	9.3	744	14	R0377154	602001018	591	19	9.3	864	13	R1698013	6020303310
519	9.3	745	9	AA016764	102023.5	592	19	9.3	865	13	R1156149	6020903310	
520	19	9.3	745	9	AA016764	102023.5	593	19	9.3	867	13	R0962333	602027134
521	19	9.3	748	13	BM617682		594	19	9.3	868	14	R1639662	602084511
522	19	9.3	748	13	R1411866	602048549	595	19	9.3	869	13	AI575876	A1575876
523	19	9.3	749	9	AI061778	1034887.5	596	19	9.3	869	13	AI140465	6020963123
524	19	9.3	751	9	AA949995	1029587.5	597	19	9.3	872	14	R0722238	A00722238
525	19	9.3	754	9	AI575876	A1575876	598	19	9.3	875	12	R0343266	6020430457
526	19	9.3	754	14	R1969586	603147792	599	19	9.3	877	13	R0972308	6020841364
527	19	9.3	757	13	R1694533	603341913	600	19	9.3	879	14	R0670984	A00670984
528	19	9.3	758	14	AI575921	A1575921	601	19	9.3	879	14	R0646564	A00646564
529	19	9.3	759	14	R1144989	602010529	602	19	9.3	882	12	R013459	601760668
530	19	9.3	760	14	R1555640	603239680	603	19	9.3	883	14	R0719173	A00719173
531	19	9.3	760	14	092014	602013	604	19	9.3	884	17	08034551	A0034551
532	19	9.3	761	9	AA951433	10341868.5	605	19	9.3	884	14	R0233102	A00233102
533	19	9.3	761	12	R0575354	602597830	606	19	9.3	887	12	R0164110	601773389
534	19	9.3	761	12	R0123419	601759115	607	19					

649	19	9.3	984	13	B1649763	603276690	722	18	8.8	524	14	BQ580502	BQ580502
650	19	9.3	985	12	BQ298889	602397375	723	18	8.8	524	14	BQ580520	BQ580520
651	19	9.3	985	13	B1414345	602990359	724	18	8.8	526	12	BQ146450	BQ146450
652	19	9.3	990	12	BQ205901	601869515	725	18	8.8	526	13	BQ456230	BQ456230
653	19	9.3	991	9	AL524214	AL524214	726	18	8.8	540	9	AL154467	AL154467
654	19	9.3	991	14	BQ918349	AGENCOURT	727	18	8.8	540	14	BQ580517	BQ580517
655	19	9.3	992	14	BQ689327	AGENCOURT	728	18	8.8	540	14	BQ580517	BQ580517
656	19	9.3	993	13	BQ463421	AGENCOURT	729	18	8.8	540	13	BQ580517	BQ580517
657	19	9.3	993	14	BQ463421	AGENCOURT	730	18	8.8	540	14	BQ580517	BQ580517
658	19	9.3	996	9	A0091195	A0091195	731	18	8.8	547	9	A0091195	A0091195
659	19	9.3	996	14	BQ585594	AGENCOURT	732	18	8.8	547	14	BQ585594	BQ585594
660	19	9.3	1011	14	BQ585594	AGENCOURT	733	18	8.8	547	14	BQ585594	BQ585594
661	19	9.3	1014	12	BQ789647	601481404	734	18	8.8	547	12	BQ789647	BQ789647
662	19	9.3	1017	13	BQ549877	AGENCOURT	735	18	8.8	547	13	BQ549877	BQ549877
663	19	9.3	1031	13	BQ554691	AGENCOURT	736	18	8.8	548	13	BQ554691	BQ554691
664	19	9.3	1047	13	B1414574	AGENCOURT	737	18	8.8	551	14	BQ554691	BQ554691
665	19	9.3	1058	14	BQ423529	AGENCOURT	738	18	8.8	552	14	BQ423529	BQ423529
666	19	9.3	1060	9	A0091016	A0091016	739	18	8.8	552	9	A0091016	A0091016
667	19	9.3	1061	13	BQ554691	AGENCOURT	740	18	8.8	557	13	BQ554691	BQ554691
668	19	9.3	1063	13	BQ554691	AGENCOURT	741	18	8.8	557	13	BQ554691	BQ554691
669	19	9.3	1076	13	BQ544498	AGENCOURT	742	18	8.8	562	13	BQ544498	BQ544498
670	19	9.3	1076	14	BQ223335	AGENCOURT	743	18	8.8	562	14	BQ223335	BQ223335
671	19	9.3	1077	14	BQ223335	AGENCOURT	744	18	8.8	569	14	BQ223335	BQ223335
672	19	9.3	1081	14	BQ223335	AGENCOURT	745	18	8.8	579	14	BQ223335	BQ223335
673	19	9.3	1093	13	BQ547917	AGENCOURT	746	18	8.8	582	13	BQ547917	BQ547917
674	19	9.3	1093	13	BQ547917	AGENCOURT	747	18	8.8	582	13	BQ547917	BQ547917
675	19	9.3	1093	13	BQ547917	AGENCOURT	748	18	8.8	582	13	BQ547917	BQ547917
676	19	9.3	1093	13	BQ547917	AGENCOURT	749	18	8.8	582	13	BQ547917	BQ547917
677	19	9.3	1104	10	BQ250475	AGENCOURT	750	18	8.8	582	10	BQ250475	BQ250475
678	19	9.3	1119	13	BQ547917	AGENCOURT	751	18	8.8	582	13	BQ547917	BQ547917
679	19	9.3	1120	14	BQ547917	AGENCOURT	752	18	8.8	582	14	BQ547917	BQ547917
680	19	9.3	1126	14	BQ547917	AGENCOURT	753	18	8.8	582	14	BQ547917	BQ547917
681	19	9.3	1136	13	BQ547917	AGENCOURT	754	18	8.8	582	13	BQ547917	BQ547917
682	19	9.3	1142	12	BQ547917	AGENCOURT	755	18	8.8	582	12	BQ547917	BQ547917
683	19	9.3	1160	11	BQ547917	AGENCOURT	756	18	8.8	582	11	BQ547917	BQ547917
684	19	9.3	1169	13	BQ547917	AGENCOURT	757	18	8.8	582	13	BQ547917	BQ547917
685	19	9.3	1169	13	BQ547917	AGENCOURT	758	18	8.8	582	13	BQ547917	BQ547917
686	19	9.3	1172	14	BQ547917	AGENCOURT	759	18	8.8	582	14	BQ547917	BQ547917
687	19	9.3	1201	13	BQ547917	AGENCOURT	760	18	8.8	582	13	BQ547917	BQ547917
688	19	9.3	1213	13	BQ547917	AGENCOURT	761	18	8.8	582	13	BQ547917	BQ547917
689	19	9.3	1218	13	BQ547917	AGENCOURT	762	18	8.8	582	13	BQ547917	BQ547917
690	19	9.3	1222	13	BQ547917	AGENCOURT	763	18	8.8	582	13	BQ547917	BQ547917
691	19	9.3	1334	13	BQ547917	AGENCOURT	764	18	8.8	582	13	BQ547917	BQ547917
692	19	9.3	1334	13	BQ547917	AGENCOURT	765	18	8.8	582	13	BQ547917	BQ547917
693	19	9.3	1532	13	BQ547917	AGENCOURT	766	18	8.8	582	13	BQ547917	BQ547917
694	19	9.3	148	9	A0214791	A0214791	767	18	8.8	582	9	A0214791	A0214791
695	18	8.8	196	9	AA364995	AA364995	768	18	8.8	582	9	AA364995	AA364995
696	18	8.8	226	13	B1322553	B1322553	769	18	8.8	582	13	B1322553	B1322553
697	18	8.8	311	10	BQ735002	BQ735002	770	18	8.8	582	10	BQ735002	BQ735002
698	18	8.8	313	10	BQ735002	BQ735002	771	18	8.8	582	10	BQ735002	BQ735002
699	18	8.8	321	10	BQ735002	BQ735002	772	18	8.8	582	10	BQ735002	BQ735002
700	18	8.8	377	14	BQ735002	BQ735002	773	18	8.8	582	14	BQ735002	BQ735002
701	18	8.8	423	13	B1593878	B1593878	774	18	8.8	582	13	B1593878	B1593878
702	18	8.8	427	9	AA993440	AA993440	775	18	8.8	582	9	AA993440	AA993440
703	18	8.8	427	14	BQ745362	BQ745362	776	18	8.8	582	14	BQ745362	BQ745362
704	18	8.8	442	12	BQ235748	BQ235748	777	18	8.8	582	12	BQ235748	BQ235748
705	18	8.8	472	13	BQ516507	BQ516507	778	18	8.8	582	13	BQ516507	BQ516507
706	18	8.8	479	13	BQ516507	BQ516507	779	18	8.8	582	13	BQ516507	BQ516507
707	18	8.8	484	13	BQ516507	BQ516507	780	18	8.8	582	13	BQ516507	BQ516507
708	18	8.8	484	13	BQ516507	BQ516507	781	18	8.8	582	13	BQ516507	BQ516507
709	18	8.8	485	13	BQ516507	BQ516507	782	18	8.8	582	13	BQ516507	BQ516507
710	18	8.8	494	14	BQ626743	BQ626743	783	18	8.8	582	14	BQ626743	BQ626743
711	18	8.8	494	14	BQ626743	BQ626743	784	18	8.8	582	14	BQ626743	BQ626743
712	18	8.8	495	13	BQ626743	BQ626743	785	18	8.8	582	13	BQ626743	BQ626743
713	18	8.8	503	10	AA700542	AA700542	786	18	8.8	582	10	AA700542	AA700542
714	18	8.8	504	9	A1520141	A1520141	787	18	8.8	582	9	A1520141	A1520141
715	18	8.8	507	9	A1325715	A1325715	788	18	8.8	582	9	A1325715	A1325715
716	18	8.8	510	13	B1611092	B1611092	789	18	8.8	582	13	B1611092	B1611092
717	18	8.8	511	13	B1581486	B1581486	790	18	8.8	582	13	B1581486	B1581486
718	18	8.8	512	14	BQ666141	BQ666141	791	18	8.8	582	14	BQ666141	BQ666141
719	18	8.8	518	14	BQ666141	BQ666141	792	18	8.8	582	14	BQ666141	BQ666141
720	18	8.8	524	13	BQ580437	BQ580437	793	18	8.8	582	13	BQ580437	BQ580437
721	18	8.8	524	14	BQ580437	BQ580437	794	18	8.8	582	14	BQ580437	BQ580437


```

795 18 8-8 923 9 A0207528 A0207528 862 16 7-8 614 9 A037749 862 16 7-8 614 9 A037749
796 18 8-8 944 14 B0689256 AGENC08RT 869 16 7-8 618 9 A062454 869 16 7-8 618 9 A062454
797 18 8-8 958 9 A0220867 870 16 7-8 624 9 A0108493 870 16 7-8 624 9 A0108493
798 17 8-3 249 9 AA455551 871 16 7-8 632 10 AA114070 871 16 7-8 632 10 AA114070
799 17 8-3 251 9 AA405482 872 16 7-8 638 9 AA017869 872 16 7-8 638 9 AA017869
800 17 8-3 320 13 B1175095 B018008A4 873 16 7-8 646 9 A1148312 873 16 7-8 646 9 A1148312
801 17 8-3 508 12 B0119944 B01758026 874 16 7-8 646 9 A1657496 874 16 7-8 646 9 A1657496
802 17 8-3 508 13 B1588259 B019409.5 875 16 7-8 653 9 A137512 875 16 7-8 653 9 A137512
803 17 8-3 525 12 B0725732 B014607.7 876 16 7-8 653 9 A137512 876 16 7-8 653 9 A137512
804 17 8-3 532 14 B0824501 K-0310098 877 16 7-8 670 10 B0641071 877 16 7-8 670 10 B0641071
805 17 8-3 540 13 B1462095 B045555.1 878 16 7-8 674 13 B0650777 878 16 7-8 674 13 B0650777
806 17 8-3 540 14 B0763817 K-EST0045 879 16 7-8 680 10 B0469938 879 16 7-8 680 10 B0469938
807 17 8-3 543 13 B1170648 B012171.5 880 16 7-8 694 14 B0201789 880 16 7-8 694 14 B0201789
808 17 8-3 543 13 B1236433 B1236433 881 16 7-8 699 13 B0694907 881 16 7-8 699 13 B0694907
809 17 8-3 543 14 B0819569 B-EST0087 882 16 7-8 699 13 B0694907 882 16 7-8 699 13 B0694907
810 17 8-3 544 9 A1785617 B044810.7 883 16 7-8 782 12 B0217111 883 16 7-8 782 12 B0217111
811 17 8-3 544 13 B1258259 B1258259 884 16 7-8 822 14 B0230243 884 16 7-8 822 14 B0230243
812 17 8-3 549 12 B1241952 B1241952 885 16 7-8 822 14 B0230243 885 16 7-8 822 14 B0230243
813 17 8-3 549 13 B0673576 B1241952 886 16 7-8 822 14 B0230243 886 16 7-8 822 14 B0230243
814 17 8-3 574 14 B0109401 B0109401 887 16 7-8 1052 13 B1231494 887 16 7-8 1052 13 B1231494
815 17 8-3 624 13 B1476570 B1476570 888 16 7-8 1052 14 B0930107 888 16 7-8 1052 14 B0930107
816 17 8-3 624 13 A0445566 A0445566 889 16 7-8 199 10 A045432 889 16 7-8 199 10 A045432
817 17 8-3 627 14 B0207021 B0207021 890 16 7-8 279 9 A1205132 890 16 7-8 279 9 A1205132
818 17 8-3 628 13 B123494 B123494 891 16 7-8 411 14 B0372828 891 16 7-8 411 14 B0372828
819 17 8-3 638 9 A1643955 A1643955 892 16 7-8 446 14 B1847614 892 16 7-8 446 14 B1847614
820 17 8-3 643 10 A0546760 B0108012 893 16 7-8 446 14 B1847614 893 16 7-8 446 14 B1847614
821 17 8-3 653 13 B0261939 B0261939 894 16 7-8 455 13 B1476570 894 16 7-8 455 13 B1476570
822 17 8-3 716 13 B1297573 B1297573 895 16 7-8 455 13 B1476570 895 16 7-8 455 13 B1476570
823 17 8-3 722 9 A0001954 A0001954 896 16 7-8 503 13 B1555574 896 16 7-8 503 13 B1555574
824 17 8-3 849 9 A125185 A125185 897 16 7-8 510 13 B1239411 897 16 7-8 510 13 B1239411
825 16 7-8 258 12 B0814905 B0814905 898 16 7-8 511 13 B1239411 898 16 7-8 511 13 B1239411
826 16 7-8 293 10 A0844949 A0844949 899 16 7-8 519 9 AA392540 899 16 7-8 519 9 AA392540
827 16 7-8 307 9 A0615913 A0615913 900 16 7-8 521 14 B0264860 900 16 7-8 521 14 B0264860
828 16 7-8 316 14 B067345 B067345 901 16 7-8 525 14 B0264860 901 16 7-8 525 14 B0264860
829 16 7-8 364 14 B06557 B06557 902 16 7-8 530 9 A0560319 902 16 7-8 530 9 A0560319
830 16 7-8 367 14 B05405 B05405 903 16 7-8 534 13 B1239411 903 16 7-8 534 13 B1239411
831 16 7-8 370 9 A0556572 A0556572 904 16 7-8 534 13 B1239411 904 16 7-8 534 13 B1239411
832 16 7-8 379 9 A0473026 A0473026 905 16 7-8 534 13 B1239411 905 16 7-8 534 13 B1239411
833 16 7-8 385 14 B033847 B033847 906 16 7-8 534 13 B1239411 906 16 7-8 534 13 B1239411
834 16 7-8 387 14 B01700 B01700 907 16 7-8 534 13 B1239411 907 16 7-8 534 13 B1239411
835 16 7-8 390 14 B14924 B14924 908 16 7-8 534 13 B1239411 908 16 7-8 534 13 B1239411
836 16 7-8 401 14 B05130 B05130 909 16 7-8 534 13 B1239411 909 16 7-8 534 13 B1239411
837 16 7-8 412 9 A0467416 A0467416 910 16 7-8 534 13 B1239411 910 16 7-8 534 13 B1239411
838 16 7-8 427 9 A0433567 A0433567 911 16 7-8 534 13 B1239411 911 16 7-8 534 13 B1239411
839 16 7-8 427 14 B13473 B13473 912 16 7-8 534 13 B1239411 912 16 7-8 534 13 B1239411
840 16 7-8 433 14 B07003 B07003 913 16 7-8 534 13 B1239411 913 16 7-8 534 13 B1239411
841 16 7-8 436 9 A0999447 A0999447 914 16 7-8 534 13 B1239411 914 16 7-8 534 13 B1239411
842 16 7-8 453 9 A0445512 A0445512 915 16 7-8 534 13 B1239411 915 16 7-8 534 13 B1239411
843 16 7-8 454 14 B06952 B06952 916 16 7-8 534 13 B1239411 916 16 7-8 534 13 B1239411
844 16 7-8 456 9 A0682959 A0682959 917 16 7-8 534 13 B1239411 917 16 7-8 534 13 B1239411
845 16 7-8 468 9 A0209608 A0209608 918 16 7-8 534 13 B1239411 918 16 7-8 534 13 B1239411
846 16 7-8 482 9 A0419112 A0419112 919 16 7-8 534 13 B1239411 919 16 7-8 534 13 B1239411
847 16 7-8 491 9 A123236 A123236 920 16 7-8 534 13 B1239411 920 16 7-8 534 13 B1239411
848 16 7-8 499 12 B0815226 B0815226 921 16 7-8 534 13 B1239411 921 16 7-8 534 13 B1239411
849 16 7-8 500 9 A0494266 A0494266 922 16 7-8 534 13 B1239411 922 16 7-8 534 13 B1239411
850 16 7-8 509 9 A0530771 A0530771 923 16 7-8 534 13 B1239411 923 16 7-8 534 13 B1239411
851 16 7-8 510 13 B0837352 B0837352 924 16 7-8 534 13 B1239411 924 16 7-8 534 13 B1239411
852 16 7-8 511 9 A087325 A087325 925 16 7-8 534 13 B1239411 925 16 7-8 534 13 B1239411
853 16 7-8 524 9 A087325 A087325 926 16 7-8 534 13 B1239411 926 16 7-8 534 13 B1239411
854 16 7-8 534 12 B0201216 B0201216 927 16 7-8 534 13 B1239411 927 16 7-8 534 13 B1239411
855 16 7-8 549 9 A0261426 A0261426 928 16 7-8 534 13 B1239411 928 16 7-8 534 13 B1239411
856 16 7-8 553 9 A075043 A075043 929 16 7-8 534 13 B1239411 929 16 7-8 534 13 B1239411
857 16 7-8 559 9 A0276030 A0276030 930 16 7-8 534 13 B1239411 930 16 7-8 534 13 B1239411
858 16 7-8 571 9 A0269418 A0269418 931 16 7-8 534 13 B1239411 931 16 7-8 534 13 B1239411
859 16 7-8 583 9 A0596466 A0596466 932 16 7-8 534 13 B1239411 932 16 7-8 534 13 B1239411
860 16 7-8 589 9 A0232732 A0232732 933 16 7-8 534 13 B1239411 933 16 7-8 534 13 B1239411
861 16 7-8 591 9 A0239570 A0239570 934 16 7-8 534 13 B1239411 934 16 7-8 534 13 B1239411
862 16 7-8 592 9 A0239570 A0239570 935 16 7-8 534 13 B1239411 935 16 7-8 534 13 B1239411
863 16 7-8 594 9 A0239570 A0239570 936 16 7-8 534 13 B1239411 936 16 7-8 534 13 B1239411
864 16 7-8 594 9 A0239570 A0239570 937 16 7-8 534 13 B1239411 937 16 7-8 534 13 B1239411
865 16 7-8 604 9 A0154371 A0154371 938 16 7-8 534 13 B1239411 938 16 7-8 534 13 B1239411
866 16 7-8 610 10 B0468557 B0468557 939 16 7-8 534 13 B1239411 939 16 7-8 534 13 B1239411
867 16 7-8 614 9 A0000616 A0000616 940 16 7-8 534 13 B1239411 940 16 7-8 534 13 B1239411

```

Source	Organism	Accession	Length	Score	Percent Similarity	Best Local Similarity	Query Match	DB	US-09-727-770-2 (1-20%) x AA57390.3 (1-620)	QY	PH	RESULT 2	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AW020600	Human	AF020600	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020601	Human	AF020601	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020602	Human	AF020602	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020603	Human	AF020603	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020604	Human	AF020604	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020605	Human	AF020605	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020606	Human	AF020606	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020607	Human	AF020607	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020608	Human	AF020608	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020609	Human	AF020609	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020610	Human	AF020610	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020611	Human	AF020611	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020612	Human	AF020612	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	EST 12-SEP-1997
AW020613	Human	AF020613	620	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	Human	Human	EST 12-SEP-1997	

[illegible]

[illegible]


```

ACCESSION   B0334842
VERSION     B0334842.1
KEYWORDS    GI:20988353
SOURCE      human.
ORGANISM    Homo sapiens

REFERENCE   1 (bases 1 to 416)
AUTHORS    Dias Neto, E., Garcia Correia, R., Verjovski-Almeida, S., Brites, M. R.,
            Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,
            Goldman, G. H., Carvalho, A. F., Matsukuma, A., Raja, G. S., Simpson, D. H.,
            Brunstein, A., de Oliveira, P. S., Bucher, P., Jongseneel, C. V., O'Hare
            M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and
            Simpson, A. J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 4491-4496 (2000)
MEDLINE     20292663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-27094922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICP Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/seq.html?pl=11&R3&L2=BC6-MT0061-
            310700-021-D10&C3=2000-07-31&L4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 15
            High quality sequence stop: 415.
            Location/Qualifiers
                location=Homo sapiens"
                db_xref="taxon:9606"
                clone_lib="MT0061"
                dev_stage="Adult"
                note="organ: marrow; Vector: puc18; Site 1: Sma1; Site 2:
                Sma1; A mini-library was made by cloning products derived
                from QUESTES PCR (U.S. Letters Patent application No. 196
                7716 - Ludwig Institute for Cancer Research) profiles
                into the puc 18 vector. Reverse transcription of tissue
                mRNA and cDNA amplification were performed under low
                stringency conditions."
                BASE COUNT      80 a 112 c 116 g 78 t
                ORIGIN
                Alignment Scores:
                Pred. No.:      1,42e-08      Length:      416
                Score:          19.00      Matches:      19
                Percent Similarity: 100.00%      Conservative: 0
                Best local Similarity: 100.00%      Mismatches: 0
                Query Match:      9,27%      Indels:      0
                DB:              14      Gaps:      0

US-09-727-770-2 (1-205) x B0334842 (1-416)

QY 115 GlyThrAlaGlnProArgLeuPheValGlyMetHleuGluIleuAlaPheAla 133
|||||
DB 227 GGCAAGCCGACAGCCCGACATTCGGGGCATGATCTTGATCTTCGCGC 171
|||||

RESULT 24
BM826778
LOCUS      K121039848 S14K402s1 Homo sapiens cDNA clone S14K402s1-16-G11 57
DEFINITION mRNA sequence.
ACCESSION   BM826778
VERSION     BM826778.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens

Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 417)
Kim, N. S., Bahu, Y., Oh, J. H., Lee, J. Y., Ahn, H. Y., Cha, M. Y., Kim, M. K.,
Oh, K. J., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, S. C., Kim, S. and
Kim, Y. S.
21C Proutier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Poon-dong Yusong-gu, Taejeon 305-385, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: ycasu@proutier.kribb.co.kr
Plate: 16; row: G; column: 11
High quality sequence stop: 417.
Location/Qualifiers
    location=Homo sapiens"
    db_xref="taxon:9606"
    clone_lib="S14K402s1-16-G11"
    clone_lib "S14K402s1"
    clone_lib "K402"
    host "Top10"
    note="organ: Stomach; Vector: pTZ19KRI; Site 1: EcoRI;
    Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
    bacterial alkaline phosphatase (BAP) and then dephosphorylated
    with tobacco acid pyrophosphatase (TAP). The dephosphorylated
    intact mRNA was ligated with DNA RNA linker including EcoR
    I site by treatment of T4 RNA ligase and the first strand
    cDNA was synthesized from oligo dT selected mRNA by
    priming with di-tailed vector. The di-tailed vector was
    adjusted to have about 60nt. The cDNA vector was
    circularized with E. coli RNA ligase after digestion of
    EcoRI which site is also included in vector. An RNA strand
    converted to a cDNA strand by a 47-mer PCR method. The
    obtained cDNA vectors were used for transformation of the
    competent cells E. coli Top10" by electroporation method.
    The cDNA libraries constructed by this method are
    full length enriched cDNA library. After analyzing and
    sequencing about 2,000 - 4,000 colonies in original cDNA
    library, the abundant cDNAs were selected and amplified by
    PCR reaction using vector region primer including 72
    promoter as 5' primer and NdeI14 as 3' primer. The PCR
    products were used as template for synthesis of
    biotinylated single stranded cDNA by in vitro transcription
    reaction. The synthesized cDNA probes were hybridized with
    antisense single stranded cDNAs prepared from original
    library and incubated with avidin gel. After removing
    DNA-RNA hybrids by centrifuge, the subcloned cDNA
    libraries were constructed by transformation of the
    remaining DNA into competent cells E. coli Top10" with
    electroporation method."
    BASE COUNT      61 a 143 c 126 g 87 t
    ORIGIN
    Alignment Scores:
    Pred. No.:      1,44e-08      Length:      417
    Score:          19.00      Matches:      19
    Percent Similarity: 100.00%      Conservative: 0
    Best local Similarity: 100.00%      Mismatches: 0
    Query Match:      9,27%      Indels:      0
    DB:              14      Gaps:      0

US-09-727-770-2 (1-205) x BM826778 (1-417)

QY 115 GlyThrAlaGlnProArgLeuPheValGlyMetHleuGluIleuAlaPheAla 133
|||||
DB 320 GGCAAGCCGACAGCCCGACATTCGGGGCATGATCTTGATCTTCGCGC 171
|||||

RESULT 25
BQ372823

```



```

Query Match: 9.27% Indels: 0
DB: 14 Gaps: 0

US-09-727-770-2 (1-295) x BM619687 (1-464)

QY 115 GlyThrAlaGlnProArqLeupheValGlyMetIleLeuIleLeuIlePheAla 133
|||||
DB 178 GGATCTGGCGACAGCGGAGCTGTGGTGGGATGATCGATGATCGTCAATTGGCG 234
|||||

RESULT 36
LOCUS BE654182 465 bp mRNA linear EST 06-SEP-2000
DEFINITION U1-M-ANI-ale-c-06-0-01.11 NIH-BMAP-MBC-N Mus musculus cDNA clone
U1-M-ANI-ale-c-06-0-01 5', mRNA sequence.
ACCESSION BE654182
VERSION BE654182.1 GI:9480095
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 465)
AUTHORS Ronaldo, M.F., Lennon, G., and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery. Genome Res 6 (9), 791-806 (1996)
JOURNAL Genome Res 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd., Room 7N 7190, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mbe@mail.nih.gov
cDNA Library Preparation: M.B. Soares Lab. cDNA distribution.
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements.
Seq primer: M13 Reverse
FEATURES
Source
Location/Qualifiers
1..465
/organism="Mus musculus"
/strain="c57BL/6J"
/db_xref="taxon:10090"
/clone="U1-M-ANI-ale-c-06-0-01"
/clone_lib="NIH-BMAP-MBC-N"
/acc_start="27-72-bp"
/acc_host="pH10B (Life Technologies)"
/acc_vector="pF740-pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The
NIH-BMAP-MBC-N library is a normalized library constructed
from mouse basal ganglia. The tag is a string of 5
nucleotides present between the Not 1 site and the
oligo-dT track. The library was constructed as described
by Ronaldo, Lennon and Soares, Genome Research 6: 791-806
(1996). Tissue provided by Dr. Andre Nedevich,
Zivic-Miller Laboratories."
BASE COUNT 81 a 147 c 118 g 119 t
ORIGIN
Alignment Scores:
Prod. No.: 1,630-08 Length: 465
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservations: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.27% Indels: 0
DB: 10 Gaps: 0

US-09-727-770-2 (1-295) x BE654182 (1-465)

QY 115 GlyThrAlaGlnProArqLeupheValGlyMetIleLeuIleLeuIlePheAla 133
|||||
DB 114 GGACGCGCGACGACGCGGAGCTATTGCGCGGATGATCGATGATCGTCAATTGGCG 169
|||||

RESULT 48
LOCUS BE601869 474 bp mRNA linear EST 25-APR-2001
DEFINITION BE601869 MARC 480V Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE601869
VERSION BE601869.1 GI:11695092
KEYWORDS EST.
SOURCE Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


Tel: 402 762 4366
 Fax: 402 762 4360
 Email: smithemall.mar.usda.usda.gov
 Single pass sequencing. Bases called and all trimmed with phred v0.980/04.e. Vector identified by cross_match with the minscore 18 and -minmatch 12 options.

PCR primers
 FORWARD: AGCAACACGATACACAT
 BACKWARD: GTTTCACGATACACAT
 Plate: 79 row: K column: 12
 Seq primer: ATTAGGTCATACATAG.
 Location/Qualifiers
 1..489
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MAGE 3RD"
 /issue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 Longissimus muscle."
 74 a 189 c 134 q 92 1

Alignment Scores:
 Pred. No.: 1.74e-08 Length: 489
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Mismatches: 0
 Best local similarity: 100.00% Mismatches: 0
 Query Match: 9.27% Indels: 0
 Gaps: 12

US-09-727-770-2 (1-205) x PE796928 (1-489)

QY 115 GlyTrAlaGlnGlnProArgLeuPheValGlyMetIleleuIleleuIleAla 133
 DB 112 GGCACAGCCAGCAGCGCGGCTCTCTGGTCATGATGTCATCTCTGCGC 168

RESULT 41
 BE276276
 LOCUS
 DEFINITION
 BE276276 mRNA sequence.
 BE276276.1 GI:951239
 BE276276 EST.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 491)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Contact: Robert Strausberg, Ph.D.
 Email: eqgaps@mail.nih.gov
 Tissue Procurement: ATCC/ATCC/ATCC
 cDNA Library Preparation: Ligo/Pharmacia
 cDNA Library Arrayed by: the I.M.A.C.E. Consortium (LML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.C.E. Consortium/LML at: image.llnl.gov
 Plate: LLC95 row: b column: 21
 High quality sequence stop: 489.
 Location/Qualifiers
 1..491
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone "IMAGE:950036"
 /issue_type="full length"
 /lab_host="DH10B (phage-resistant)"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..491
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone "IMAGE:950036"
 /issue_type="full length"
 /lab_host="DH10B (phage-resistant)"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..491
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone "IMAGE:950036"
 /issue_type="full length"
 /lab_host="DH10B (phage-resistant)"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..491
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone "IMAGE:950036"
 /issue_type="full length"
 /lab_host="DH10B (phage-resistant)"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..491
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone "IMAGE:950036"
 /issue_type="full length"
 /lab_host="DH10B (phage-resistant)"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..491
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone "IMAGE:950036"
 /issue_type="full length"
 /lab_host="DH10B (phage-resistant)"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

/note "Organism: skin; Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 cDNA made by oligo dT priming. Birectionally
 cloned into EcoRI/XbaI sites using the following 5'
 adaptor: GGTACAGAG(G). Site selected 500bp for average
 insert size 1.8kb. Library constructed by Ligo Bond II
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 74 a 176 c 147 q 95 1

Alignment Scores:
 Pred. No.: 1.75e-08 Length: 491
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Mismatches: 0
 Best local similarity: 100.00% Mismatches: 0
 Query Match: 9.27% Indels: 0
 Gaps: 10

US-09-727-770-2 (1-205) x PE796276 (1-491)

QY 115 GlyTrAlaGlnGlnProArgLeuPheValGlyMetIleleuIleleuIleAla 133
 DB 412 GGCACAGCCAGCAGCGCGGCTCTCTGGTCATGATGTCATCTCTGCGC 468

RESULT 42
 AW768111
 LOCUS
 DEFINITION
 AW768111 456 bp mRNA linear EST 16 FEB 2001
 IMAGE:320051 Harland stage 19 23 Xenopus laevis cDNA clone
 IMAGE:320051.1 similar to qv.Mc2/ba.VA006666 A1.157; IMAGE:31 20
 PROTEIN:110 SUBUNIT (HUMAN). Q5.M64239 Mouse vacuolar H(+) ATPase
 (MOUSE); mRNA sequence.
 AW768111
 AW768111.1 GI:7700194
 AW768111 EST.
 SOURCE
 ORGANISM
 African clawed frog.
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 Xenopodidae; Xenopus.
 1 (bases 1 to 456)
 Clifton, S., Johnson, S., L., Blumberg, B., Som, J., Hillier, L., Papadopoulos,
 Martin, J., Wyllie, J., Underwood, K., Theis, B., Bowers, V., Person,
 B., Gibbons, M., Harvey, N., Ketter, J., Jackson, J., McCann, K.,
 Waterston, R., and Wilson, R.
 WashU Xenopus EST project, 1999
 Unpublished (1999)
 Contact: Sandy Clifton, Ph.D.
 WashU Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8601, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Library constructed by R. Harland, Ph.D. (University of California,
 Berkeley)

FEATURES
 SOURCE
 Location/Qualifiers
 1..456
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone "IMAGE:320051"
 /issue_type="Harland stage 19 23"
 /lab_host="stage 19 23"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..456
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone "IMAGE:320051"
 /issue_type="Harland stage 19 23"
 /lab_host="stage 19 23"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..456
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone "IMAGE:320051"
 /issue_type="Harland stage 19 23"
 /lab_host="stage 19 23"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..456
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone "IMAGE:320051"
 /issue_type="Harland stage 19 23"
 /lab_host="stage 19 23"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..456
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone "IMAGE:320051"
 /issue_type="Harland stage 19 23"
 /lab_host="stage 19 23"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..456
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone "IMAGE:320051"
 /issue_type="Harland stage 19 23"
 /lab_host="stage 19 23"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

FEATURES
 SOURCE
 Location/Qualifiers
 1..456
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone "IMAGE:320051"
 /issue_type="Harland stage 19 23"
 /lab_host="stage 19 23"
 /note "Vector: pCMV sport6, site_1: NotI; Site_2: SalI;
 Library constructed by"

Tel: 414 286 1800
 Fax: 414 286 1810
 Email: estwatson.wustl.edu
 Library constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B. Gurdon, (Wellcome/3Kc Institute). DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/BLND at: info@image.tul.tgo
 Seq primer: -40RP from Gibco
 High quality sequence stop: 459.

FEATURES

source

1..512
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:343578"
 /clone_lib="Wellcome CRC p30NA1 eqq"
 /issue_type="eqq"
 /lab_host="DH10B (phage resistant)"
 /note="Vector, p30NA1, Site 1, Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Library was constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute)."
 116 a 127 c 108 q 161 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 1,84c-08 Length: 512
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.27% Indels: 0
 DB: 12 Gaps: 0

US-09 727-770-2 (1-205) x BG445676 (1-512)

QY 115 GYThrAlaGlnProArgGluGluValGlyMetIleLeuIleLeuIlePheAla 133
 |||||||
 DB 209 GGIACAGCAGACAAACCTGGACTATCTGGGAATGATCTGATCTGATCTGCA 265

RESULT 50

AL644755

LOCUS

AL644755 XCC-cgg Silurana tropicalis cDNA clone 1114q09 5', mRNA
 sequence.

ACCESSION

AL644755

VERSION

AL644755.1 GI:16796880

KEYWORDS

EST

SOURCE

western clawed frog.

ORGANISM

Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Silurana.

1 (bases 1 to 513)

Huckle, E., Taylor, P., Ashurst, I. L., Zorn, A.M., and Rogers, J.

Sanjour Xenopus tropicalis EST project 2001 (10_2001)

Unpublished (2001)

Contact: Buckle E

Sanger Centre

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanjour Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: 1114q09.sp6

Sequencing primer: Sp6

This sequence is from a Xenopus Gene Collection (XCC) library

constructed by Aaron M. Zorn.

Location/Qualifiers

1..513

/organism="Silurana tropicalis"

/db_xref="taxon:8364"

/clone="1114q09"

/clone_lib="XCC-est"

/dev_stage="eqq"

/lab_host="Escherichia coli XL1-blue"

FEATURES

source

/note="Vector: p30NA1, Site 1, Site 2: EcoRI, Site 2: NotI; cDNA was oligo-dT primed from 5' end of poly A+ RNA from cyp. EcoRI-NotI cut cDNA was then ligated into p30NA1 with EcoRI at the 5' end and NotI at the 3' end"

BASE COUNT 92 a 144 c 149 q 148 t
 ORIGIN

Alignment Scores:

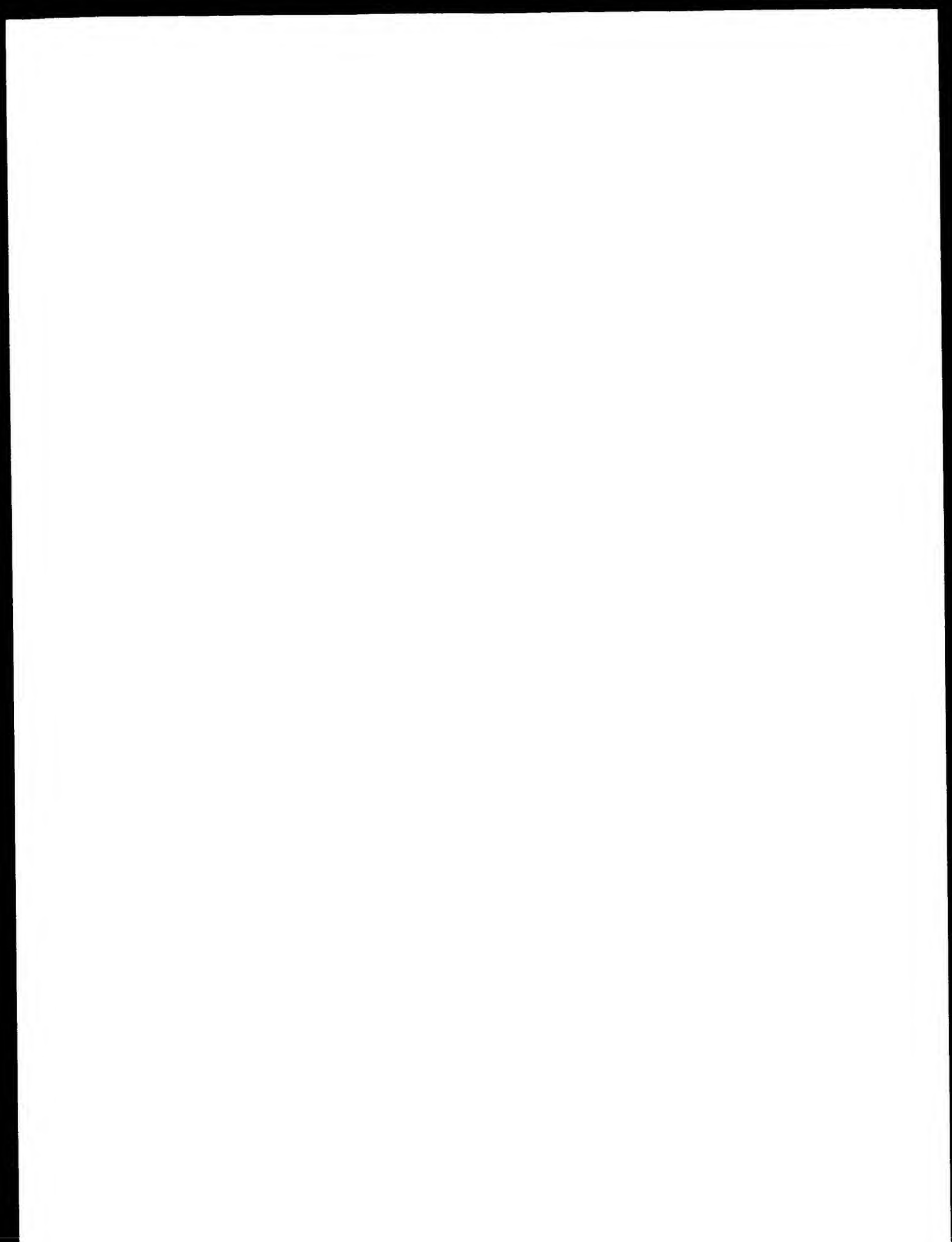
Pred. No.: 1,84c-08 Length: 513
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.27% Indels: 0
 DB: 9 Gaps: 0

US-09 727-770-2 (1-205) x AL644755 (1-513)

QY 115 GYThrAlaGlnProArgGluGluValGlyMetIleLeuIleLeuIlePheAla 133
 |||||||
 DB 432 GGIACAGCAGACAAACCTGGACTATCTGGGAATGATCTGATCTGATCTGCA 488

Search completed: December 28, 2002, 14:43:40

Job time : 2006 secs




```

PRIOR APPLICATION NUMBER: US 09/098,409
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 4917
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 15890
LENGTH: 612
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC095972.2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.9
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 15
OTHER INFORMATION: EXPRESSED IN HEK293, SIGNAL = 5.1
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 4.4
OTHER INFORMATION: NT HIT: 34562312, EVALUO 9.05E-39
OTHER INFORMATION: EST HUMAN HIT: BF526960.1, EVALUO 1.00E-38
OTHER INFORMATION: SWISSPROT hit: P25356, EVALUO 1.00E-29
US-09-864-761-16890
Alignment Scores:
Prod. No.: 9-59e-196 Length: 612
Score: 203.00 Matches: 203
Percent Similarity: 100.00 Mismatches: 0
Best Local Similarity: 100.00 Indels: 0
Query Match: 99.129 Gaps: 0
DB:
US 09-727 770 2 (1 200) x US-09-864-761-16890 (1-612)
QY 1 MetSerAsnSerProGluTerAlaLeuValPheThrIleSerGlyAlaMetAlaThr 20
|||||
DB 3 ATGTCAAAACACACCCCACTATCTTCTCTTTTAACTATCTCTCTCTCTCTCTCTCT 62
QY 21 MetValSerSerGlyLeuAlaAlaCysGlyMetAlaLysAsnGlyThrGlyMet 40
|||||
DB 63 ATGTTCTTCAAGTGGCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 122
QY 41 AlaMetSerValMetIleProGluGluIleHisMetLysSerAlaIleValValMet 60
|||||
DB 123 GCATGTCGTGCAAGGCGACAGACGATGACATGAAGGATATATATATATATATATATAT 182
QY 61 AlaGlyIleIleThrIleTyrGlyLeuValAlaAlaValProProAlaAsnSerLeuAsn 80
|||||
DB 183 GTTGGTATCATCACCACATCATGATGATGATGATGATGATGATGATGATGATGATG 242
QY 61 AspAspSerLeuTyrSerSerPheLeuIleLeuGlyAlaGlyLeuSerGlyLeuAla 100
|||||
DB 243 GATGCAAAAGAGTCGTATAGCAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 302
QY 101 AlaGlyPheAlaIleValIleValIleAspSerGlyGlySerGlyThrAlaGlyGlyPro 120
|||||
DB 303 GAGGGGTTTGGCCATGTCATGTCGAGGAGACATGGTANGTCTGACATGTCAGACACAC 362
QY 121 ArgLeuPheValGlyMetIleLeuIleLeuIlePheAlaLysValIleIleLeuSerThr 140
|||||
DB 363 CGACTATTCTAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 422
QY 141 GAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 423 AAGTAGGCTCTCAAAACCCCAATGACAGAAATAGATGATGATGATGATGATGATGAT 482
QY 161 PheArgAsnLysGlnProAspThrHisValIleGlySerThrProSerValAlaAspLeu 180
|||||
DB 483 TTGCGAAACAAACAGGTTACACAGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 542

```

```

QY 181 LeuSerValCysGlySerIleGluValHisAsnAlaGluAlaGlyPheLeuPhePro 200
|||||
DB 543 CTAAGTGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
QY 201 HisAlaVal 203
|||||
DB 603 CATGCTGTC 611
RESULT 4
US-09-864-761-48
: Sequence 48, Application US/09864761
: Patent No. US2002048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Bank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENE DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: ELITE REFERENCE: Acemich-X-1
: CURRENT APPLICATION NUMBER: US 09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 09/192,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 09/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/622,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 09/236,459
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 4917
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 48
: LENGTH: 448
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC095972.2
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.9
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.9
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.7
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.6
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4

```

```

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 15
OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL - 5.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL - 6.4
US-09-864-761-48
Alignment Scores:
Pred. No.: 2,548-116 Length: 448
Score: 124.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.49% Indels: 0
DB: 10 Gaps: 0

US-09-727-770-2 (1-205) x US-09-964-761-48 (1-448)
QY 1 MetSerAsnSerProGluTyrAlaLeuValPheThrIleSorGlyAlaMetAtaThr 20
DB 76 ATGTCCAAACAAACAGCGCGAGTATGCTTTGCTTTCACCAATCGGGTGTATGCGCCAC 135
QY 21 MetValSerSerGlyLeuGlyAlaAlaCysAlMetAlaGlyCAserGlyThrGlyIleMet 40
DB 146 ATGATCTCCAGTGGCGTGGGGTGGCGCTGTGGATGGGCAAGAAAGGACCGGCATCATG 195
QY 41 AlaMetSerValMetTrpProGluLeuIleHisMetIlySerIleIleProValValMet 60
DB 196 GGCATGTCGTGATGATGCGCCAGAGCTGATCCACATGAAGTCCATCCACAGTGGTCAATG 255
QY 61 AlaGlyIleIleGlyThrIleTyrClyLeuValAlaAlaValProProAlaAsaSerLeuAsn 80
DB 256 GCTGGTATCATCATCAGCATCATATGAGGCTAGTGGCGCTGCGCGCTGCGCACTGGCTGAAT 315
QY 81 AsnAsnAsnSerLeuTyrSerSerPheLeuGlyLeuGlyAlaGlyLeuGlyLeuAla 100
DB 416 GATGACAAACAGTCTCTATAGCAGTTTCTCCAGCTGGCGCTGCGCTGAGTGGCTGCA 375
QY 101 AlaGlyPheAlaIleValIleValIleValIleValIleValIleValIleValIleValIle 120
DB 476 GCGGCTTTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
QY 121 ArctopheVal 124
DB 436 GGACTATTGTGA 447

RESULT 5
US-09-925-401-49
Sequence 49, Application US/09925401
Patent No. US200200952408A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
FILE REFERENCE: Nucleic Acids, Proteins and Antibodies
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05482
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 69/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 49
LENGTH: 956
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47)
OTHER INFORMATION: n equals a.t.g. or c
NAME/KEY: misc_feature
LOCATION: (352)
OTHER INFORMATION: n equals a.t.g. or c
US-09-925-401-49
Alignment Scores:
Pred. No.: 2,924-08 Length: 956
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.29% Indels: 0
DB: 10 Gaps: 0

US-09-727-770-2 (1-205) x US-09-925-401-49 (1-956)
QY 117 AlaGlnProAlaLeuPheValGlyMetIleLeuIleLeuIlePheAla 144
DB 521 GCGCAGCAGCGCGAGTATGCTGGGCAATGCTGATGCTGATGCTGATGCTGATGCTGCTGCG 571

RESULT 6
US-09-887-576-814
Sequence 814, Application US/09887576
Patent No. US20020144047A1
GENERAL INFORMATION:
APPLICANT: Budworth, P.
APPLICANT: Brown, D.
APPLICANT: Chang, H.
APPLICANT: Zhu, T.
APPLICANT: Han, B.
APPLICANT: Wang, X.
APPLICANT: Cooper, Bret
TITLE OF INVENTION: Promoters for regulation of plant expression
FILE REFERENCE: 1360.001US1
CURRENT APPLICATION NUMBER: US/09/887,576
CURRENT FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: US 60/213,848
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/214,087
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/258,662
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 875
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 814
LENGTH: 498
TYPE: DNA
ORGANISM: Oryza sativa
US-09-887-576-814
Alignment Scores:
Pred. No.: 0.00171 Length: 498
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.85% Indels: 0
DB: 10 Gaps: 0

US-09-727-770-2 (1-205) x US-09-887-576-814 (1-498)
QY 122 IopPheValGlyMetIleLeuIleLeuIlePheAla 144
DB 488 GTTTTGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424

RESULT 7
US-09-887-576-815
Sequence 815, Application US/09887576
Patent No. US20020144047A1
GENERAL INFORMATION:
APPLICANT: Budworth, P.
APPLICANT: Brown, D.
APPLICANT: Chang, H.
APPLICANT: Zhu, T.
APPLICANT: Han, B.
APPLICANT: Wang, X.
APPLICANT: Cooper, Bret
TITLE OF INVENTION: Promoters for regulation of plant expression
FILE REFERENCE: 1360.001US1
CURRENT APPLICATION NUMBER: US/09/887,576
CURRENT FILING DATE: 2001-06-25
```

```

: PRIOR APPLICATION NUMBER: US 60/214,848
: PRIOR FILING DATE: 2000-06-24
: PRIOR APPLICATION NUMBER: US 60/214,087
: PRIOR FILING DATE: 2000-06-24
: PRIOR APPLICATION NUMBER: US 60/258,692
: PRIOR FILING DATE: 2000-12-29
: NUMBER OF SEQ ID NOS: 875
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 815
: LENGTH: 498
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-887-576-815

```

```

Alignment Scores:
Pred. No.: 0.00172      Length: 498
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservat: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 5.85%             Indels: 0
DB: 10                    Gaps: 0

```

US 09 727 770-2 (1-205) x US 09 887-576-815 (1-498)

```

QY 122 leupheValglyMetIleLeuIleLeuIlePheAla 133
|||||
DB 388 CTGTTGGTGGTAATGATCTTCATCTTCATTTTTCG 423

```

RESULT 8

```

US-09-887-576-775
: Sequence 775, Application US/09887575
: Patent No. US20020144047A1
: GENERAL INFORMATION:
: APPLICANT: Budworth, P.
: APPLICANT: Brown, D.
: APPLICANT: Chang, H.
: APPLICANT: Zhu, T.
: APPLICANT: Han, B.
: APPLICANT: Wang, X.
: APPLICANT: Cooper, Bret
: TITLE OF INVENTION: Promoters for regulation of plant expression
: FILE REFERENCE: 1360.001US1
: CURRENT APPLICATION NUMBER: US 09/887,575
: CURRENT FILING DATE: 2001-06-25
: PRIOR APPLICATION NUMBER: US 60/214,087
: PRIOR FILING DATE: 2000-06-24
: PRIOR APPLICATION NUMBER: US 60/214,087
: PRIOR FILING DATE: 2000-06-24
: PRIOR APPLICATION NUMBER: US 60/258,692
: PRIOR FILING DATE: 2000-12-29
: NUMBER OF SEQ ID NOS: 875
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 775
: LENGTH: 500
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-887-576-775

```

```

Alignment Scores:
Pred. No.: 0.00172      Length: 500
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservat: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 5.85%             Indels: 0
DB: 10                    Gaps: 0

```

US 09 727 770-2 (1-205) x US 09 887-576-775 (1-500)

```

QY 122 leupheValglyMetIleLeuIleLeuIlePheAla 133
|||||
DB 391 TTGTTGGTGGTAATGATCTTCATCTTCATTTTTCG 426

```

RESULT 9

```

US-09-948-842A-645
: Sequence 645, Application US/09948642A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Krops, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: FILE REFERENCE: SCRIPT400-3
: CURRENT APPLICATION NUMBER: US 09/948,842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 645
: LENGTH: 501
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-948-842A-645

```

```

Alignment Scores:
Pred. No.: 0.00172      Length: 501
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservat: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 5.85%             Indels: 0
DB: 9                    Gaps: 0

```

US-09 727 770-2 (1-205) x US-09 948-842A 645 (1-501)

```

QY 122 leupheValglyMetIleLeuIleLeuIlePheAla 133
|||||
DB 391 CTGTTGGTGGTAATGATCTTCATCTTCATTTTTCG 426

```

RESULT 10

```

US-09-887-576-786
: Sequence 786, Application US/09887576
: Patent No. US20020144047A1
: GENERAL INFORMATION:
: APPLICANT: Budworth, P.
: APPLICANT: Brown, D.
: APPLICANT: Chang, H.
: APPLICANT: Zhu, T.
: APPLICANT: Han, B.
: APPLICANT: Wang, X.
: APPLICANT: Cooper, Bret
: TITLE OF INVENTION: Promoters for regulation of plant expression
: FILE REFERENCE: 1360.001US1
: CURRENT APPLICATION NUMBER: US/09/887,576
: CURRENT FILING DATE: 2001-06-25
: PRIOR APPLICATION NUMBER: US 60/214,848
: PRIOR FILING DATE: 2000-06-24
: PRIOR APPLICATION NUMBER: US 60/214,087
: PRIOR FILING DATE: 2000-06-24
: PRIOR APPLICATION NUMBER: US 60/258,692
: PRIOR FILING DATE: 2000-12-29
: NUMBER OF SEQ ID NOS: 875
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 786
: LENGTH: 544
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-887-576-786

```

```

Alignment Scores:
Pred. No.: 0.00185      Length: 543
Score: 12.00           Matches: 12

```

US-09-727-770-2.olil0p2n.rnpb

APPLICANT: ZHU, T.

